

Determining accurate conformational ensembles of intrinsically disordered proteins at atomic resolution

Supplementary Information

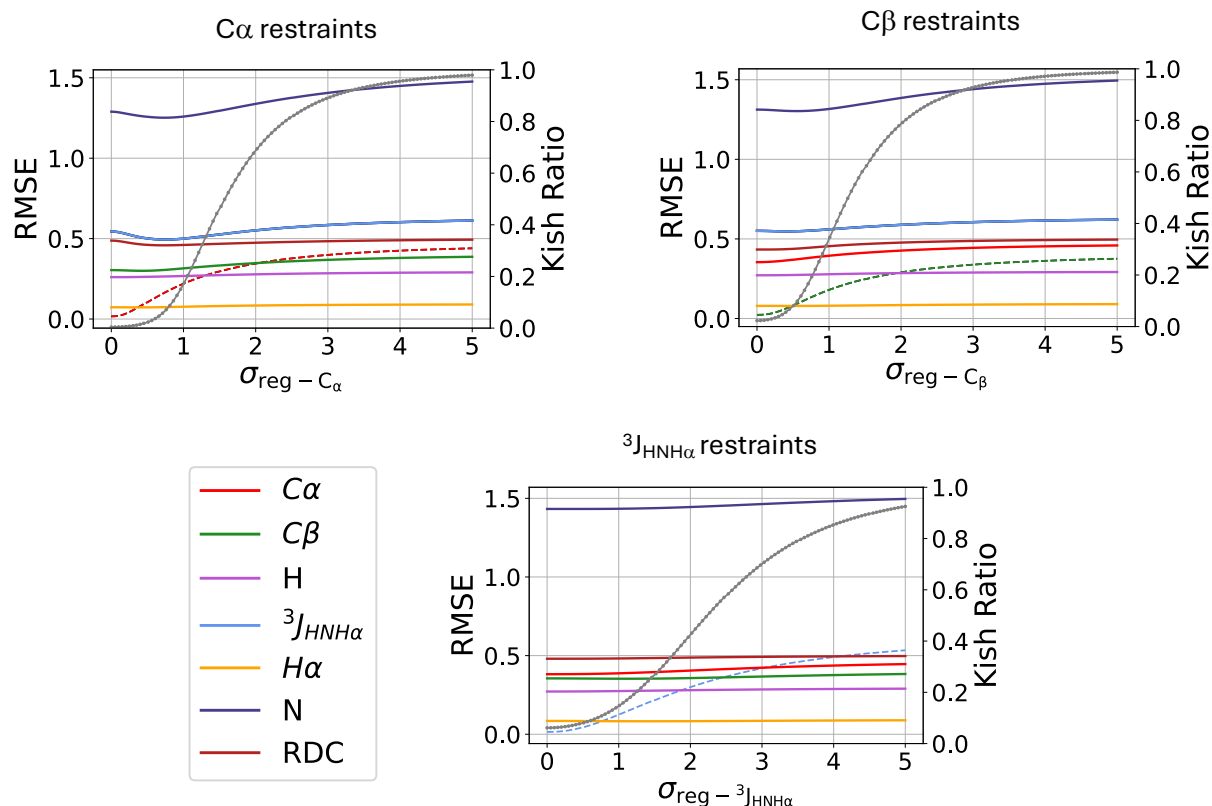
Kaushik Borthakur,¹ Thomas R. Sisk,¹ Francesco P. Panei,² Massimiliano
Bonomi,² and Paul Robustelli^{*,1}

¹*Dartmouth College, Department of Chemistry, Hanover, NH, 03755, USA*

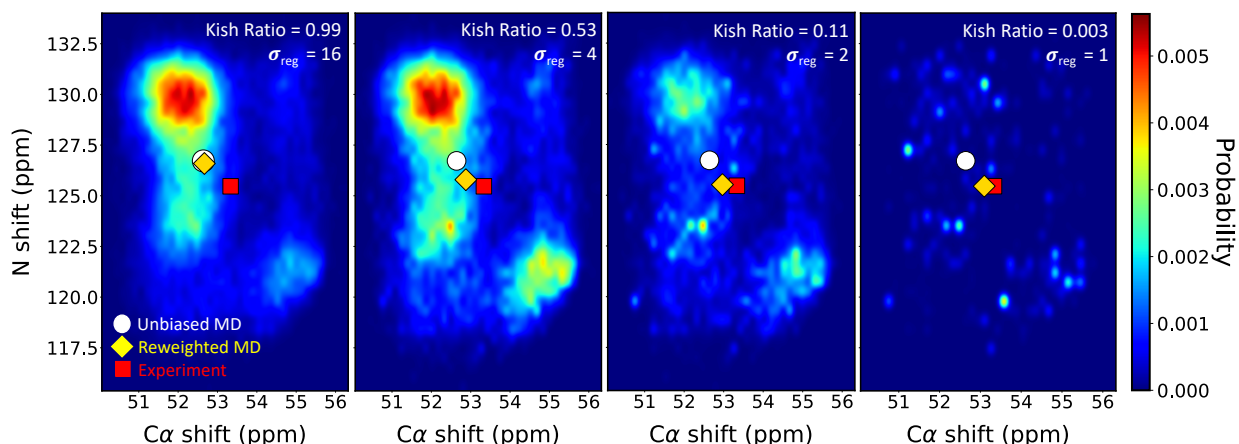
²*Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Computational Structural
Biology Unit, Paris, France*

* E-mail: paul.j.robustelli@dartmouth.edu

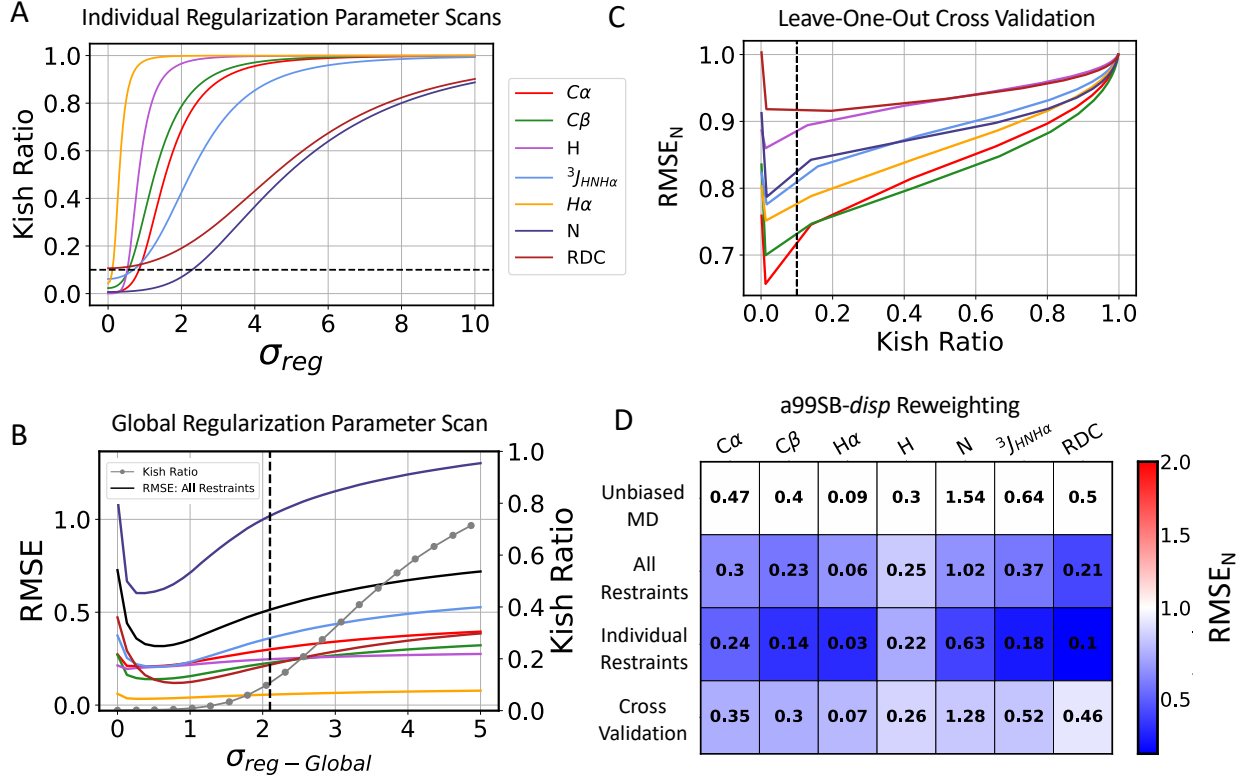
Supplementary Figures



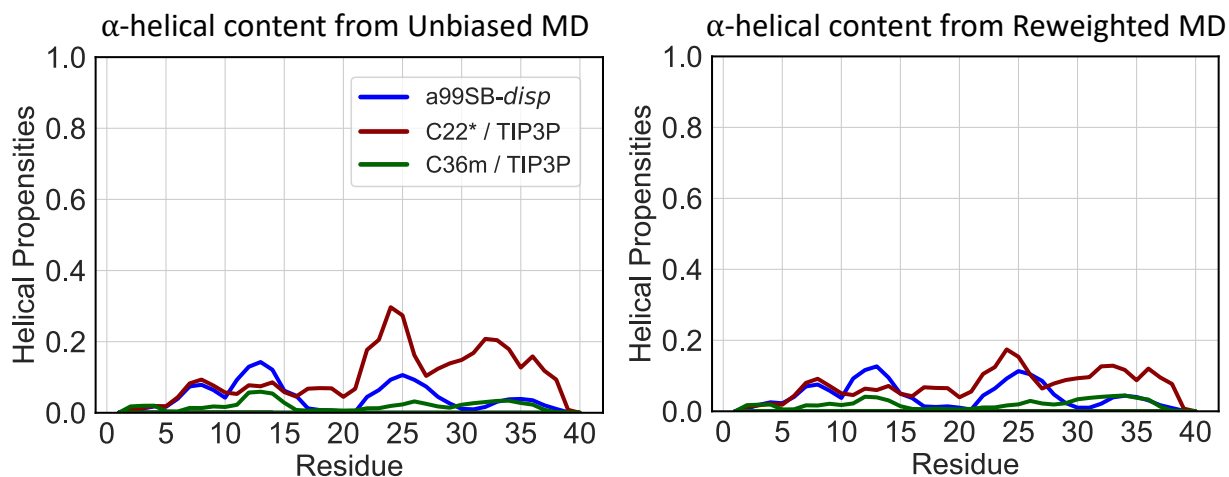
Supplementary Figure 1: Accuracy of predictions of unrestrained cross-validating data in reweighted A β 40 ensembles obtained by reweighting an a99SB-*disp* MD ensemble with different experimental restraints. The root mean square error (RMSE) between calculated and experimental data in reweighted A β 40 ensembles are shown as a function of the value of the regularization parameter (σ_{reg}) of experimental restraints used for reweighting. Experimental data types are indicated by color. The RMSEs of unrestrained cross-validating data are indicated by solid lines, and the RMSE of restrained experimental data is indicated by dotted lines. The Kish ratio of each reweighted ensemble (right y-axis) is indicated by gray dots.



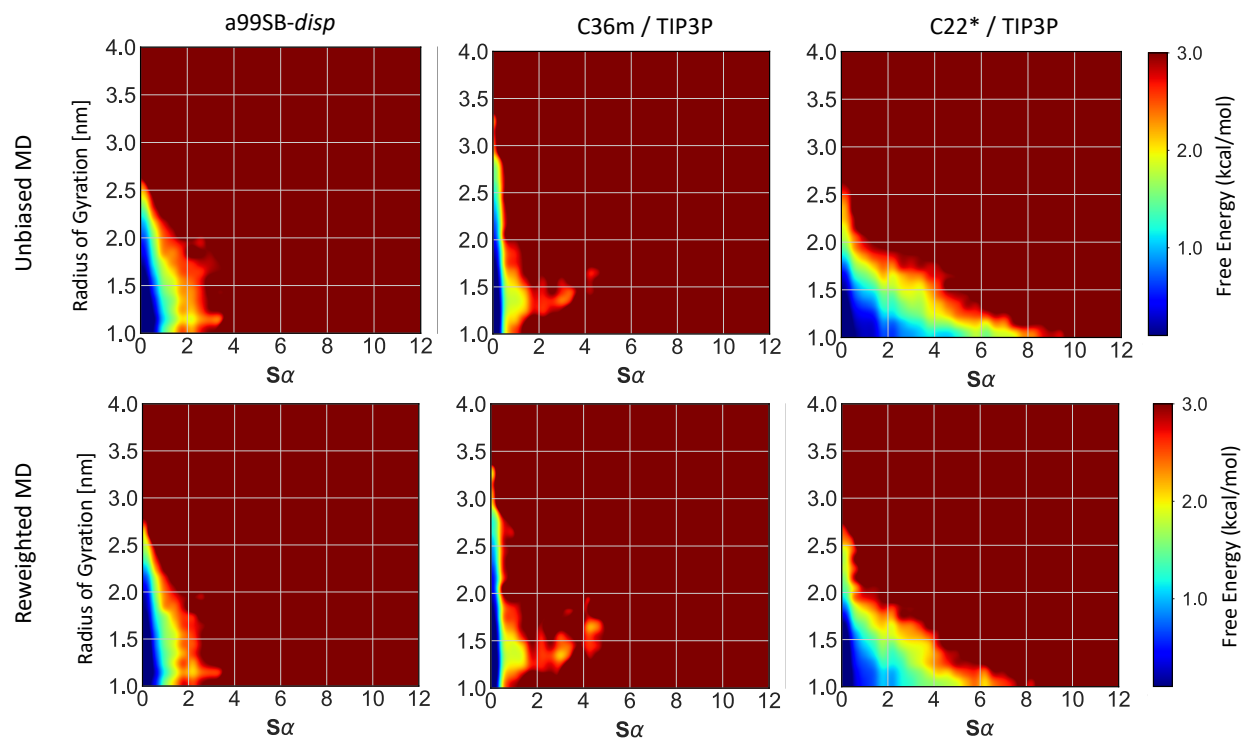
Supplementary Figure 2: Reweighted IDP ensembles become less representative of the most populated conformational states sampled in unbiased MD simulations as the strength of experimental restraints is increased. We performed maximum entropy reweighting of a $30\mu\text{s}$ unbiased a99SB-*disp* MD ensemble of the ParE2 associated antitoxin protein (PaaA2) using the backbone $C\alpha$ and N NMR chemical shifts of residue ALA38 as restraints. We compare the distribution of the predicted values of these two chemical shifts in each reweighted ensemble as a function of the value of regularization parameter σ_{reg} of these restraints. Smaller values of σ_{reg} correspond to greater strengths of experimental restraints. The calculated values of the restrained chemical shifts in each reweighted ensemble are shown as yellow diamonds. The calculated values of the chemical shifts in the unbiased MD ensemble are shown as a white circle and the experimental values of the chemical shifts are shown as a red square.



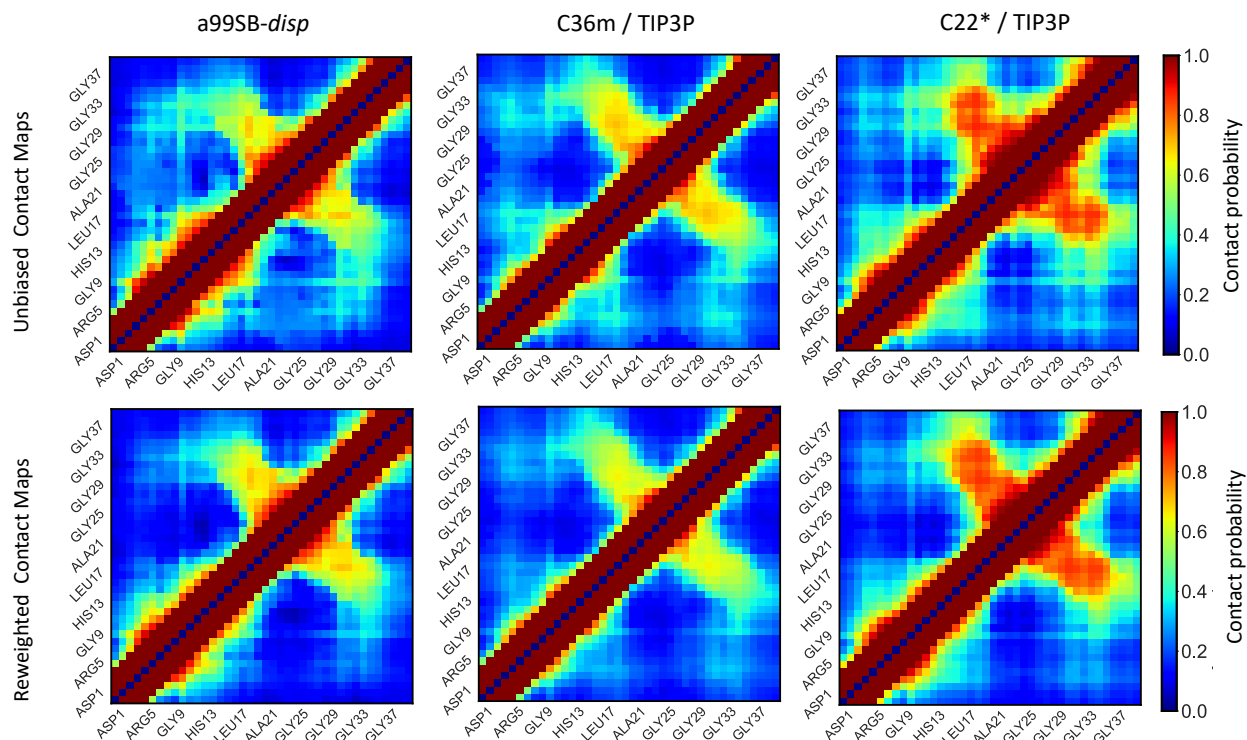
Supplementary Figure 3: Maximum entropy reweighting of an a99SB-*disp* MD ensemble of A β 40. **A)** The effective ensemble size of reweighted ensembles obtained using individual experimental data types as restraints as a function of the value of the regularization parameter σ_{reg} . **B)** Root mean square error (RMSE) between calculated and experimental data in reweighted ensembles as a function of the value of the global regularization parameter scaling factor ($\sigma_{reg-Global}$). The Kish ratio of each reweighted ensemble is indicated by gray dots. **C)** The normalized RMSE (RMSE_N) of unrestrained data in leave-one-out cross-validation tests where all other experimental data types are used as restraints for reweighting. RMSE_N values of each data type are normalized by the RMSE observed in the unbiased a99SB-*disp* A β 40 MD ensemble. Black dotted lines indicate the Kish ratio threshold $K = 0.10$ in each plot. **D)** RMSE of experimental data in reweighted ensembles obtained by reweighting the a99SB-*disp* MD ensemble with different combinations of restraints. RMSEs are displayed for the unbiased MD ensemble (“Unbiased MD”), a reweighted ensemble obtained using all experimental data as restraints (“All Restraints”), reweighted ensembles obtained using each individual data type as the only restraint (“Individual Restraints”) and reweighted ensembles obtained using all but one data type as restraints (“Cross Validation”). All reweighted ensembles were selected using a Kish ratio threshold of $K = 0.10$. Each square is colored to reflect the value of RMSE_N relative to the unbiased a99SB-*disp* MD ensemble.



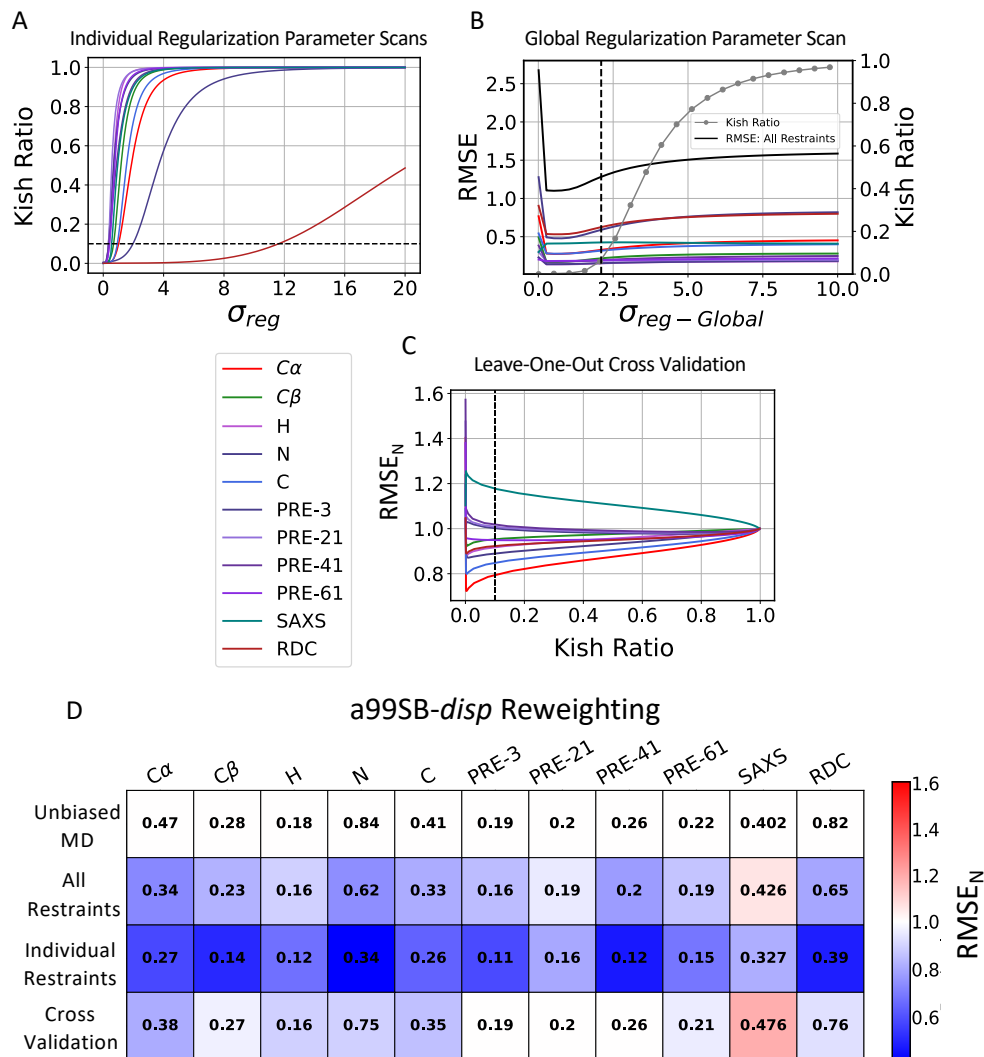
Supplementary Figure 4: Populations of α -helical conformations in unbiased and reweighted MD ensembles of A β 40.



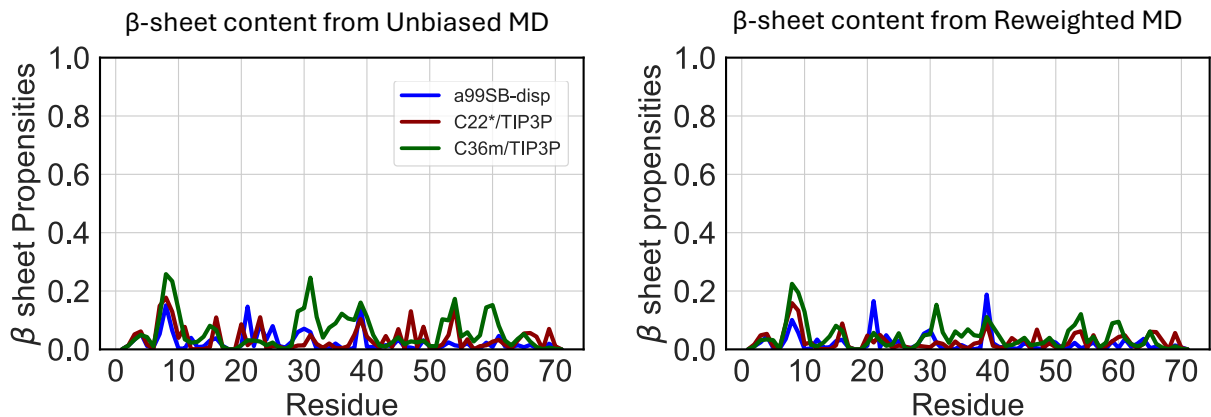
Supplementary Figure 5: Comparison of free energy surfaces of unbiased and reweighted ensembles of A β 40. Free energy surfaces of unbiased MD ensembles (top row) and reweighted ensembles (bottom row) of A β 40 are shown as a function of the radius of gyration (R_g) and α -helical order parameter S_α of each conformation.



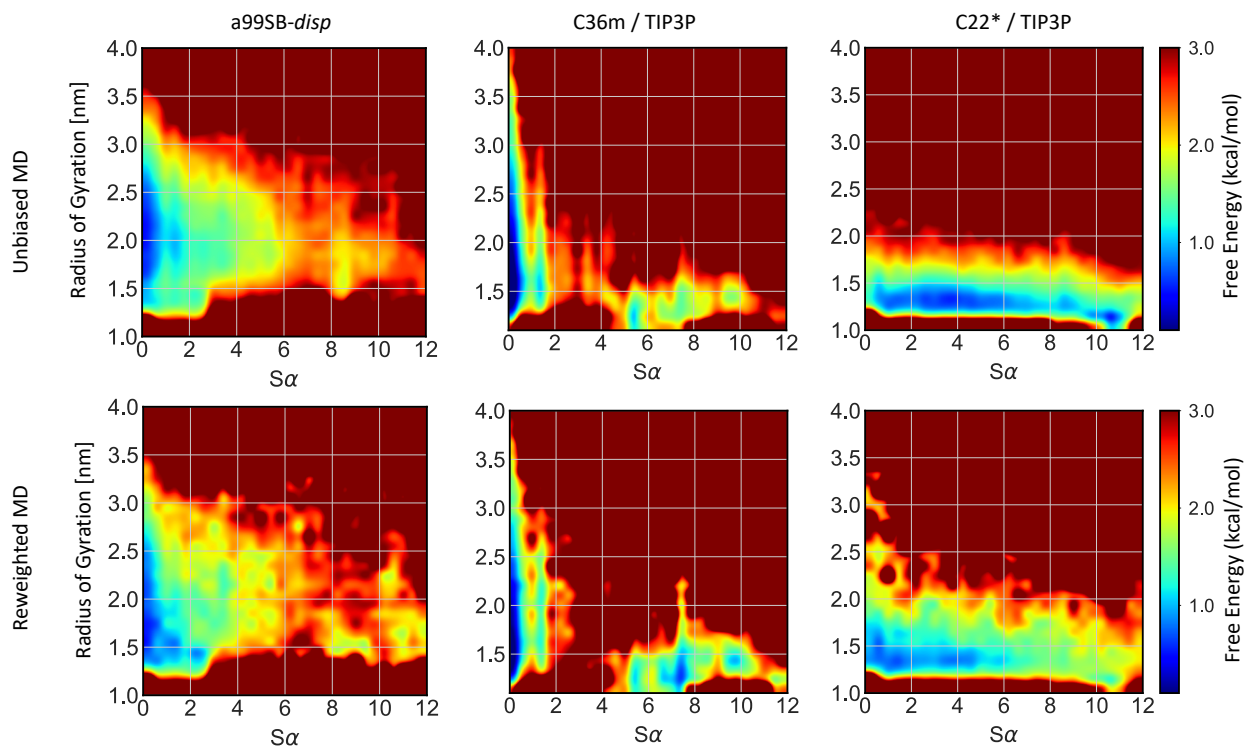
Supplementary Figure 6: Comparison of the populations of intramolecular contacts in unbiased and reweighted ensembles of A β 40. Contact maps of unbiased MD ensembles (top row) and reweighted MD ensembles (bottom row) of A β 40. Contacts between residues are defined using a 12 Å distance threshold between heavy atoms.



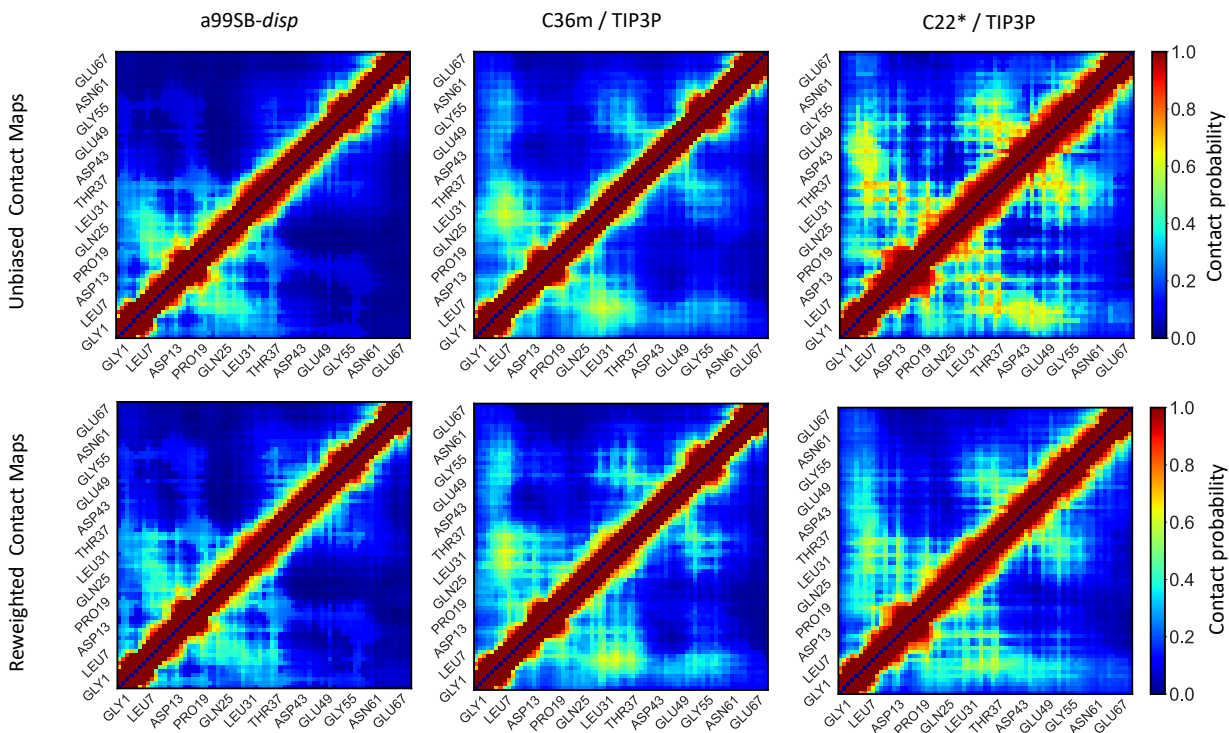
Supplementary Figure 7: Maximum entropy reweighting of an a99SB-*disp* MD ensemble of ACTR. **A)** The effective ensemble size of reweighted ensembles obtained using individual experimental data types as restraints as a function of the value of the regularization parameter σ_{reg} . **B)** Root mean square error (RMSE) between calculated and experimental data in reweighted ensembles as a function of the value of the global regularization parameter scaling factor ($\sigma_{reg-Global}$). The Kish ratio of each reweighted ensemble is indicated by gray dots. **C)** The normalized RMSE ($RMSE_N$) of unrestrained data in leave-one-out cross-validation tests where all other experimental data types are used as restraints for reweighting. $RMSE_N$ values of each data type are normalized by the RMSE observed in the unbiased a99SB-*disp* ACTR MD ensemble. Black dotted lines indicate the Kish ratio threshold $K = 0.10$ in each plot. **D)** RMSE of experimental data in reweighted ensembles obtained by reweighting the a99SB-*disp* MD ensemble with different combinations of restraints. RMSEs are displayed for the unbiased MD ensemble (“Unbiased MD”), a reweighted ensemble obtained using all experimental data as restraints (“All Restraints”), reweighted ensembles obtained using each individual data type as the only restraint (“Individual Restraints”) and reweighted ensembles obtained using all but one data type as restraints (“Cross Validation”). All reweighted ensembles were selected using a Kish ratio threshold of $K = 0.10$. Each square is colored to reflect the value of $RMSE_N$ relative to the unbiased a99SB-*disp* MD ensemble.



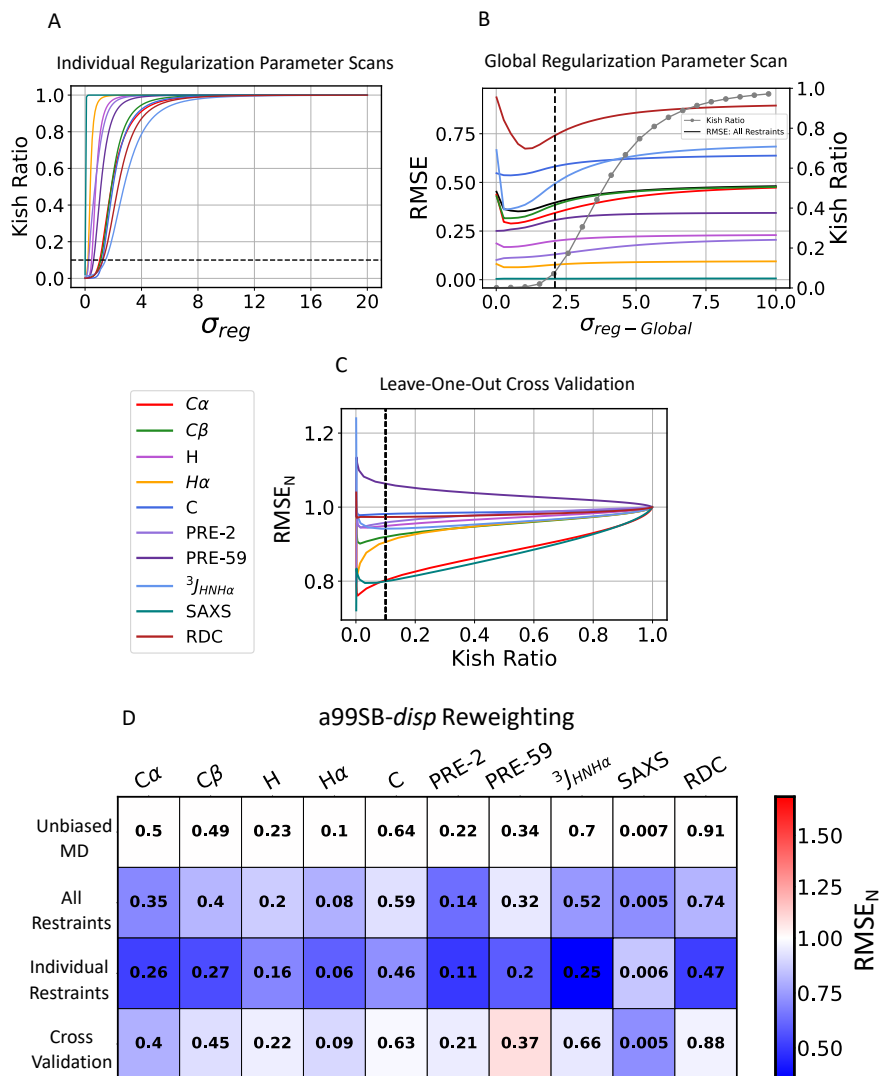
Supplementary Figure 8: Populations of β sheet conformations in unbiased and reweighted MD ensembles of ACTR.



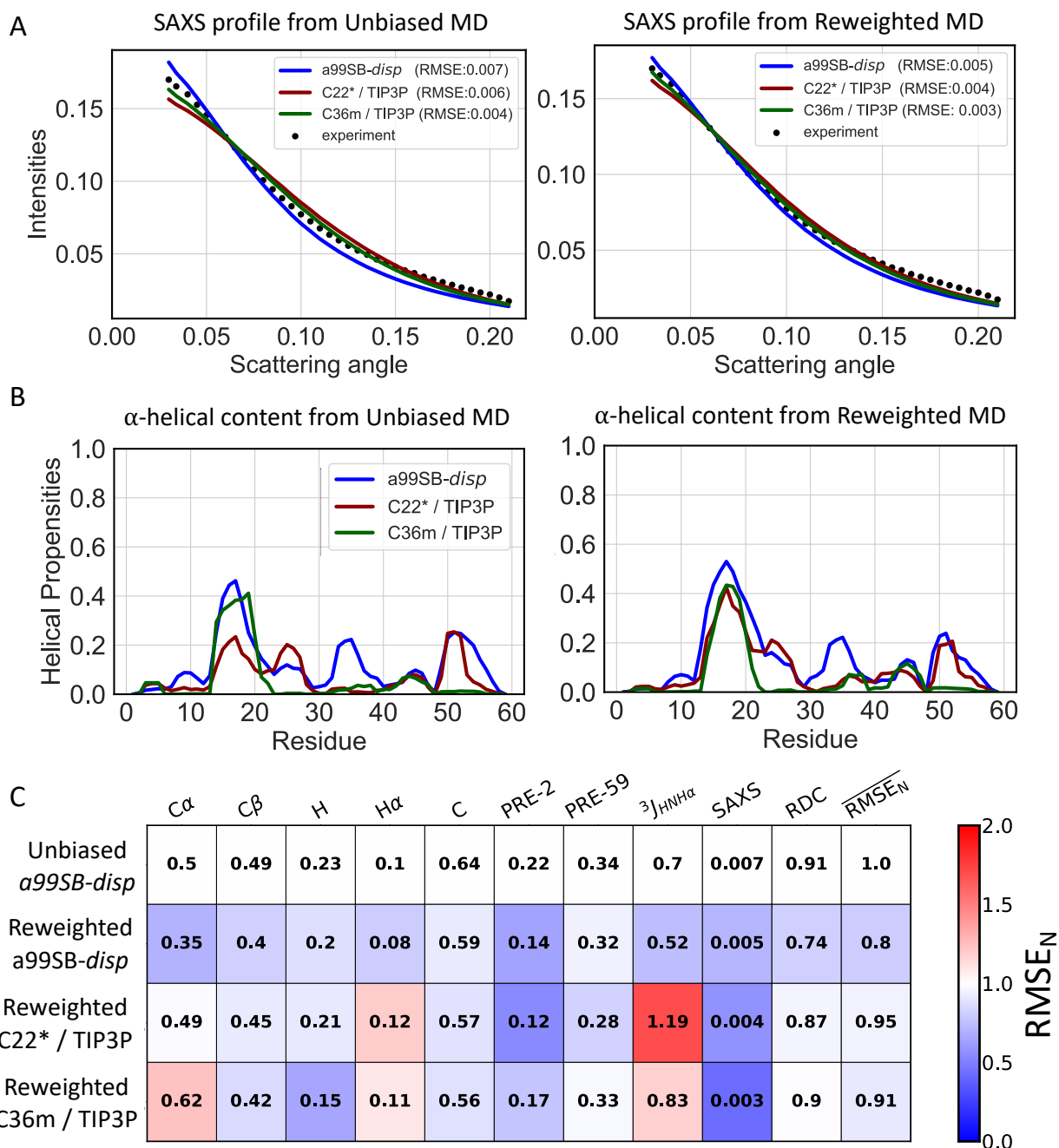
Supplementary Figure 9: Comparison of free energy surfaces of unbiased and reweighted ensembles of ACTR. Free energy surfaces of unbiased MD ensembles (top row) and reweighted ensembles (bottom row) of ACTR are shown as a function of the radius of gyration (R_g) and α -helical order parameter $S\alpha$ of each conformation.



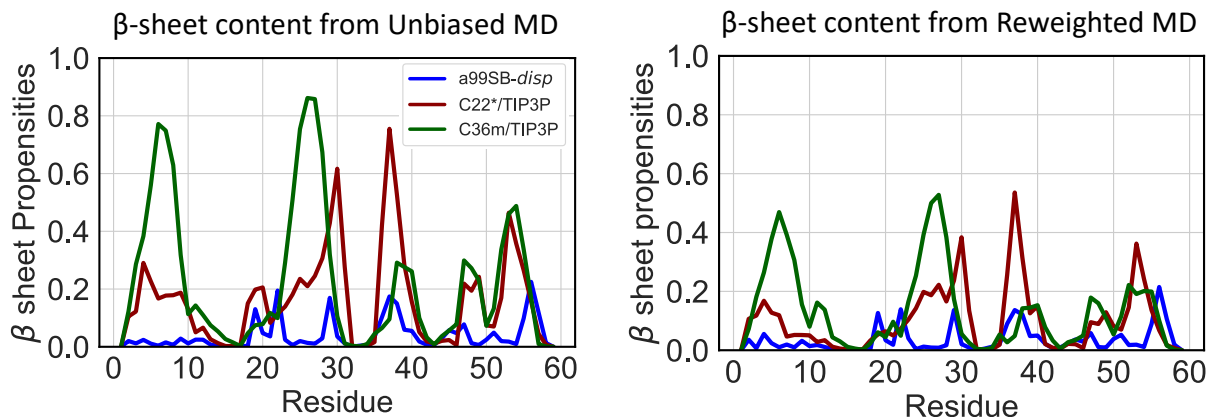
Supplementary Figure 10: Comparison of the populations of intramolecular contacts in unbiased and reweighted ensembles of ACTR. Contact maps of unbiased MD ensembles (top row) and reweighted MD ensembles (bottom row) of ACTR. Contacts between residues are defined using a 12 Å distance threshold between heavy atoms.



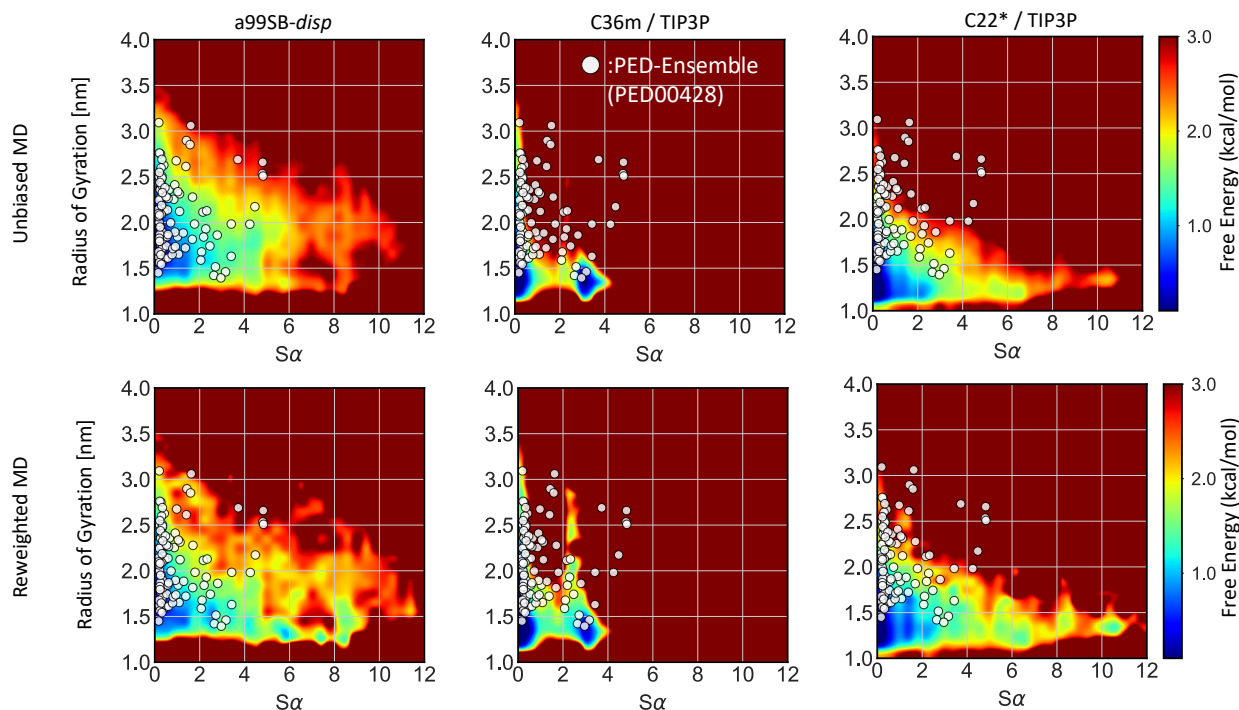
Supplementary Figure 11: Maximum entropy reweighting of an a99SB-*disp* MD ensemble of drkN SH3. **A)** The effective ensemble size of reweighted ensembles obtained using individual experimental data types as restraints as a function of the value of the regularization parameter σ_{reg} . **B)** Root mean square error (RMSE) between calculated and experimental data in reweighted ensembles as a function of the value of the global regularization parameter scaling factor ($\sigma_{reg-Global}$). The Kish ratio of each reweighted ensemble is indicated by gray dots. **C)** The normalized RMSE (RMSE_N) of unrestrained data in leave-one-out cross-validation tests where all other experimental data types are used as restraints for reweighting. RMSE_N values of each data type are normalized by the RMSE observed in the unbiased a99SB-*disp* drkN SH3 MD ensemble. Black dotted lines indicate the Kish ratio threshold $K = 0.10$ in each plot. **D)** RMSE of experimental data in reweighted ensembles obtained by reweighting the a99SB-*disp* MD ensemble with different combinations of restraints. RMSEs are displayed for the unbiased MD ensemble (“Unbiased MD”), a reweighted ensemble obtained using all experimental data as restraints (“All Restraints”), reweighted ensembles obtained using each individual data type as the only restraint (“Individual Restraints”) and reweighted ensembles obtained using all but one data type as restraints (“Cross Validation”). All reweighted ensembles were selected using a Kish ratio threshold of $K = 0.10$. Each square is colored to reflect the value of RMSE_N relative to the unbiased a99SB-*disp* MD ensemble.



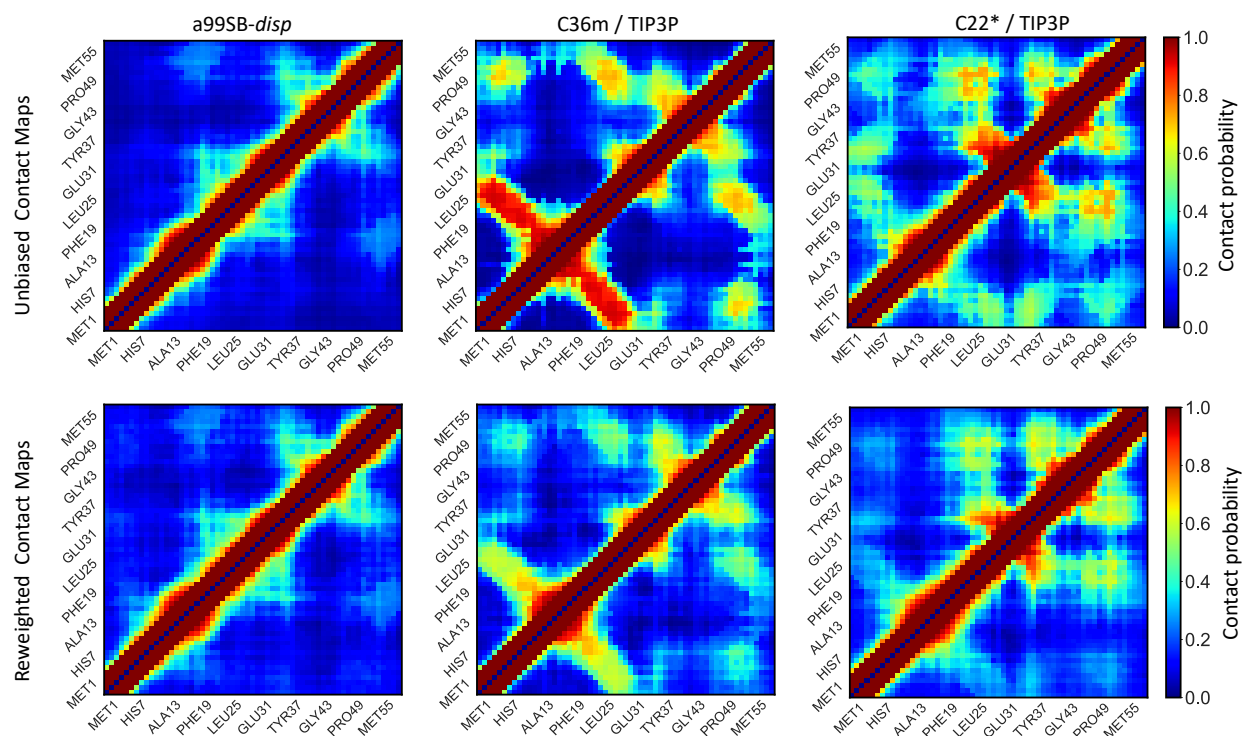
Supplementary Figure 12: Comparison of unbiased and reweighted MD ensembles of drkN SH3. **A)** Comparison of SAXS profiles calculated from unbiased MD ensembles and reweighted MD ensembles of drkN SH3. **B)** Populations of α -helical conformations in unbiased MD ensembles and reweighted MD ensembles of drkN SH3. **C)** Comparison of the RMSE between calculated and experimental data in reweighted drkN SH3 ensembles derived from different force fields. Reweighted ensembles were calculated using all experimental data as restraints with a Kish ratio threshold of $K=0.10$. Each square is colored to reflect the value of the normalized RMSE (RMSE_N) relative to the unbiased *a99SB-disp* MD ensemble.



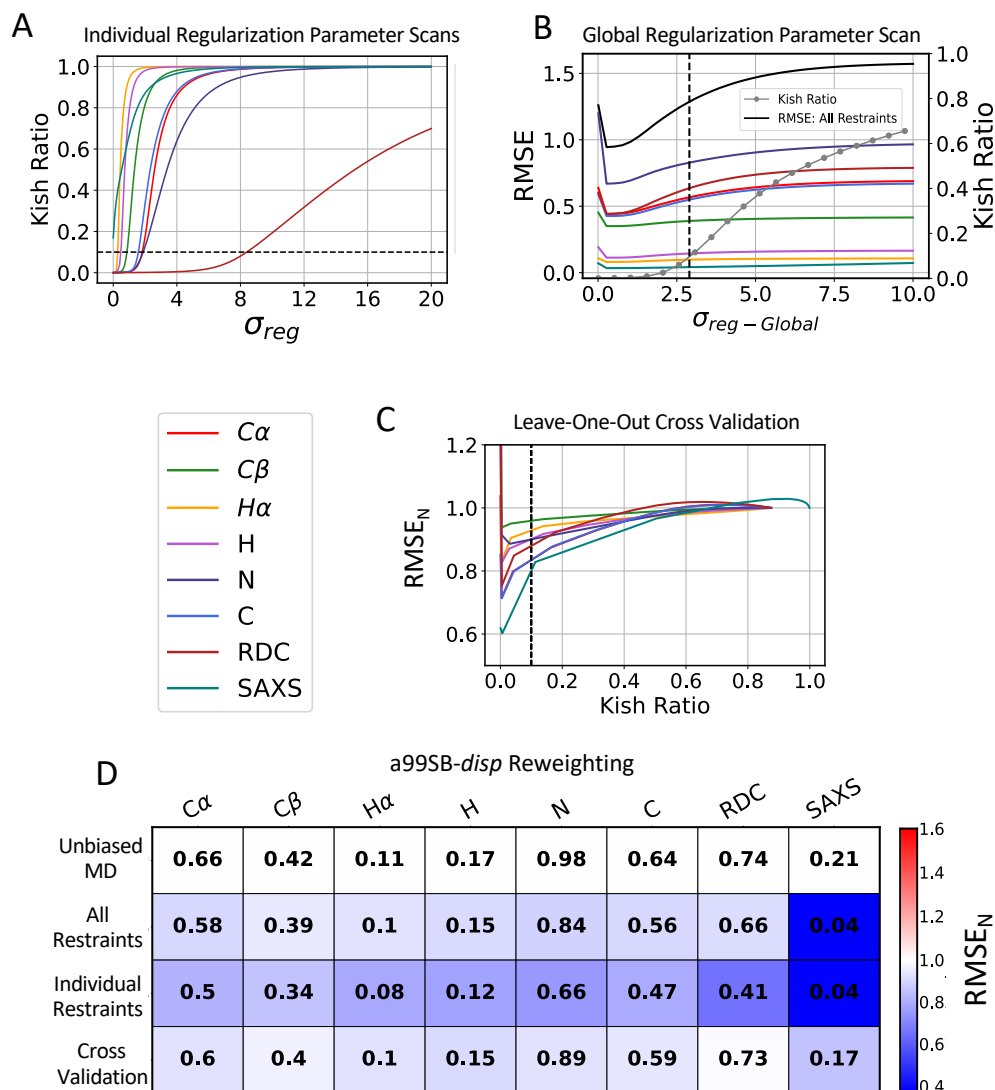
Supplementary Figure 13: Populations of β sheet conformations in unbiased and reweighted MD ensembles of drkN SH3.



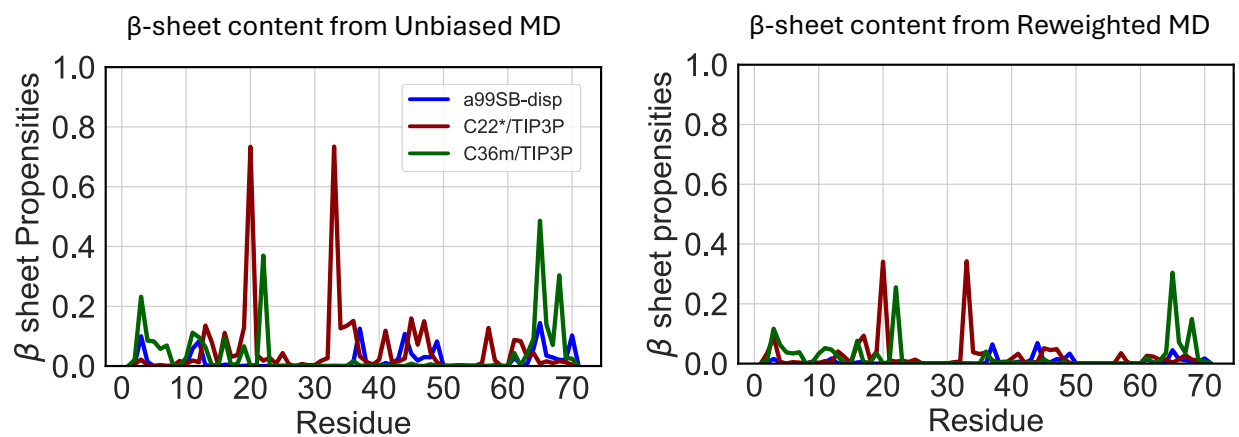
Supplementary Figure 14: Comparison of free energy surfaces of unbiased and reweighted ensembles of drkN SH3. Free energy surfaces of unbiased MD ensembles (top row) and reweighted ensembles (bottom row) of drkN SH3 are shown as a function of the radius of gyration (R_g) and α -helical order parameter $S\alpha$ of each conformation. The white dots reflect the the R_g and $S\alpha$ values of structures from a drkN SH3 protein ensemble database (PED) ensemble calculated from experimental NMR data.¹



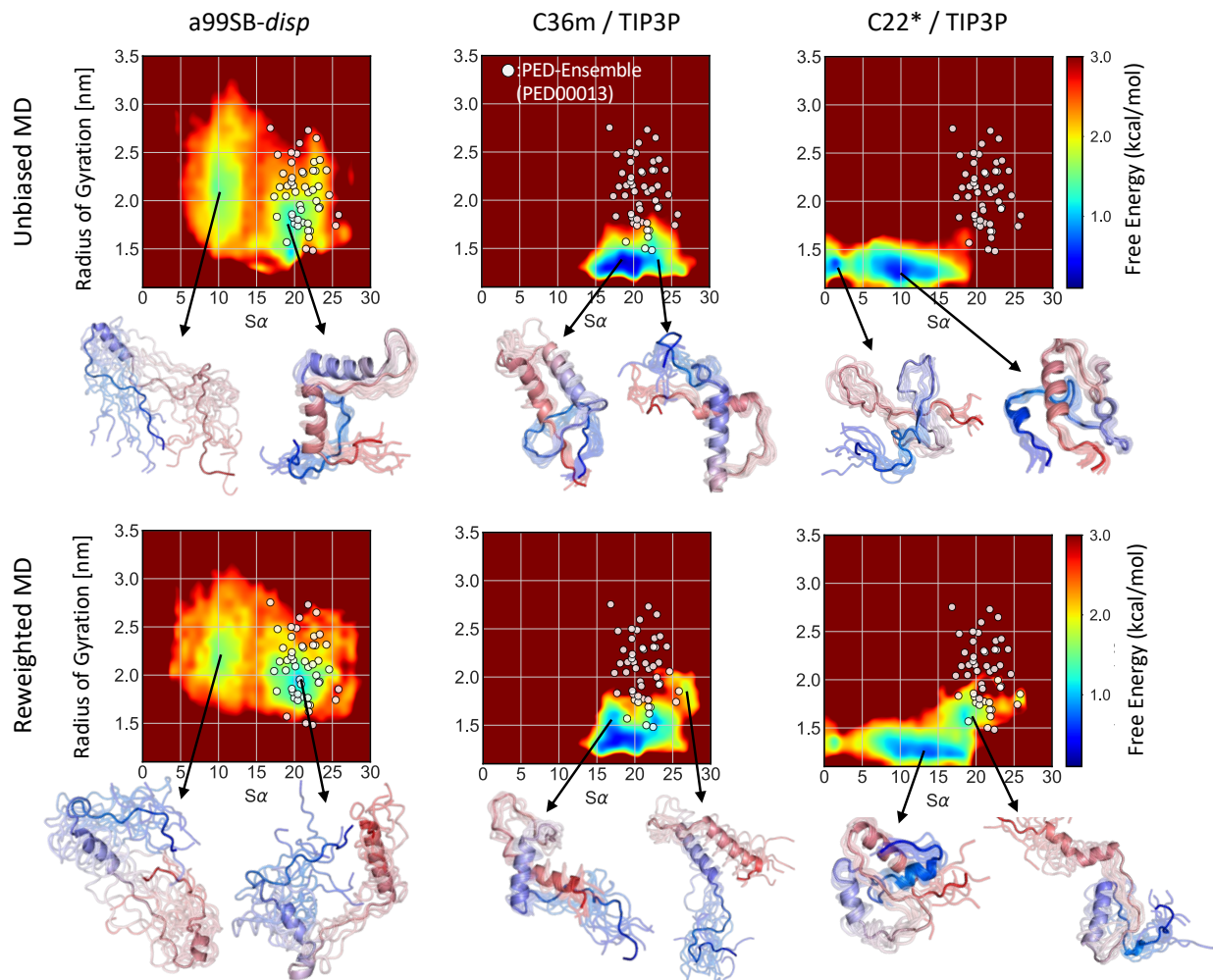
Supplementary Figure 15: Comparison of the populations of intramolecular contacts in unbiased and reweighted ensembles of drkN SH3. Contact maps of unbiased MD ensembles (top row) and reweighted MD ensembles (bottom row) of drkN SH3. Contacts between residues are defined using a 12 Å distance threshold between heavy atoms.



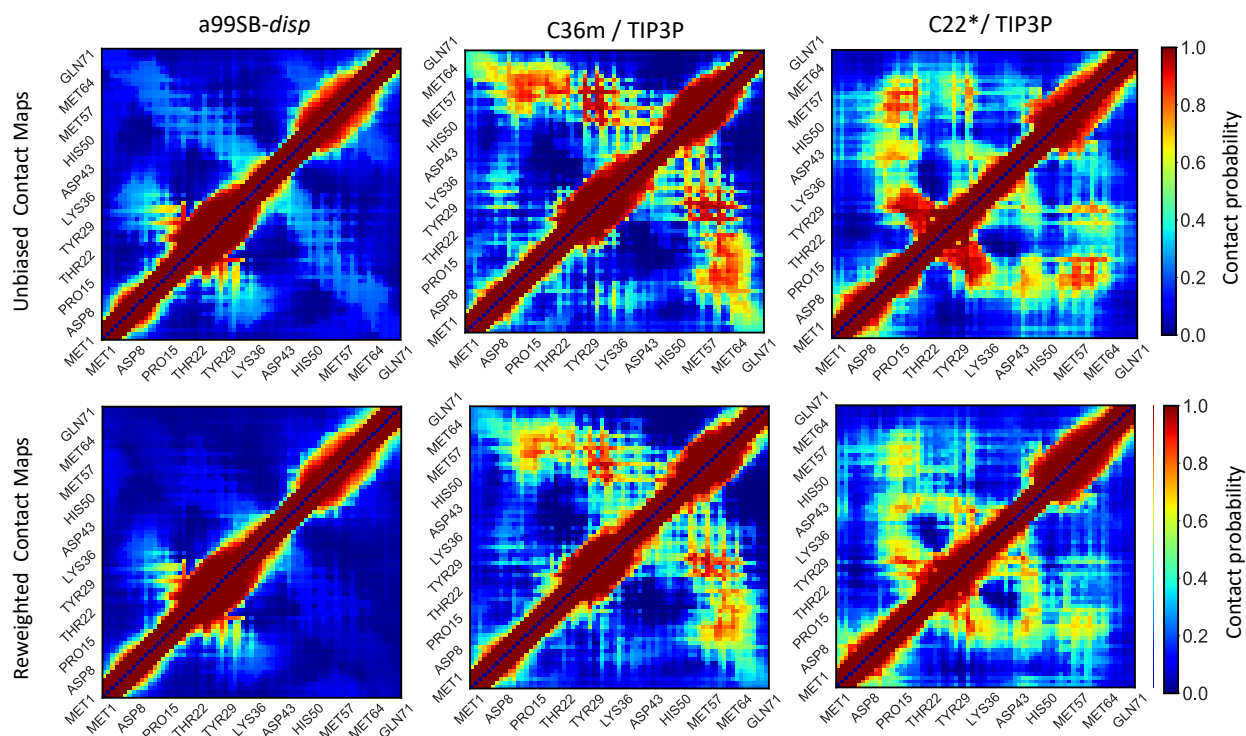
Supplementary Figure 16: Maximum entropy reweighting of an a99SB-*disp* MD ensemble of PaaA2. **A)** The effective ensemble size of reweighted ensembles obtained using individual experimental data types as restraints as a function of the value of the regularization parameter σ_{reg} . **B)** Root mean square error (RMSE) between calculated and experimental data in reweighted ensembles as a function of the value of the global regularization parameter scaling factor ($\sigma_{reg-Global}$). The Kish ratio of each reweighted ensemble is indicated by gray dots. **C)** The normalized RMSE ($RMSE_N$) of unrestrained data in leave-one-out cross-validation tests where all other experimental data types are used as restraints for reweighting. $RMSE_N$ values of each data type are normalized by the RMSE observed in the unbiased a99SB-*disp* PaaA2 MD ensemble. Black dotted lines indicate the Kish ratio threshold $K = 0.10$ in each plot. **D)** RMSE of experimental data in reweighted ensembles obtained by reweighting the a99SB-*disp* MD ensemble with different combinations of restraints. RMSEs are displayed for the unbiased MD ensemble (“Unbiased MD”), a reweighted ensemble obtained using all experimental data as restraints (“All Restraints”), reweighted ensembles obtained using each individual data type as the only restraint (“Individual Restraints”) and reweighted ensembles obtained using all but one data type as restraints (“Cross Validation”). All reweighted ensembles were selected using a Kish ratio threshold of $K = 0.10$. Each square is colored to reflect the value of $RMSE_N$ relative to the unbiased a99SB-*disp* MD ensemble.



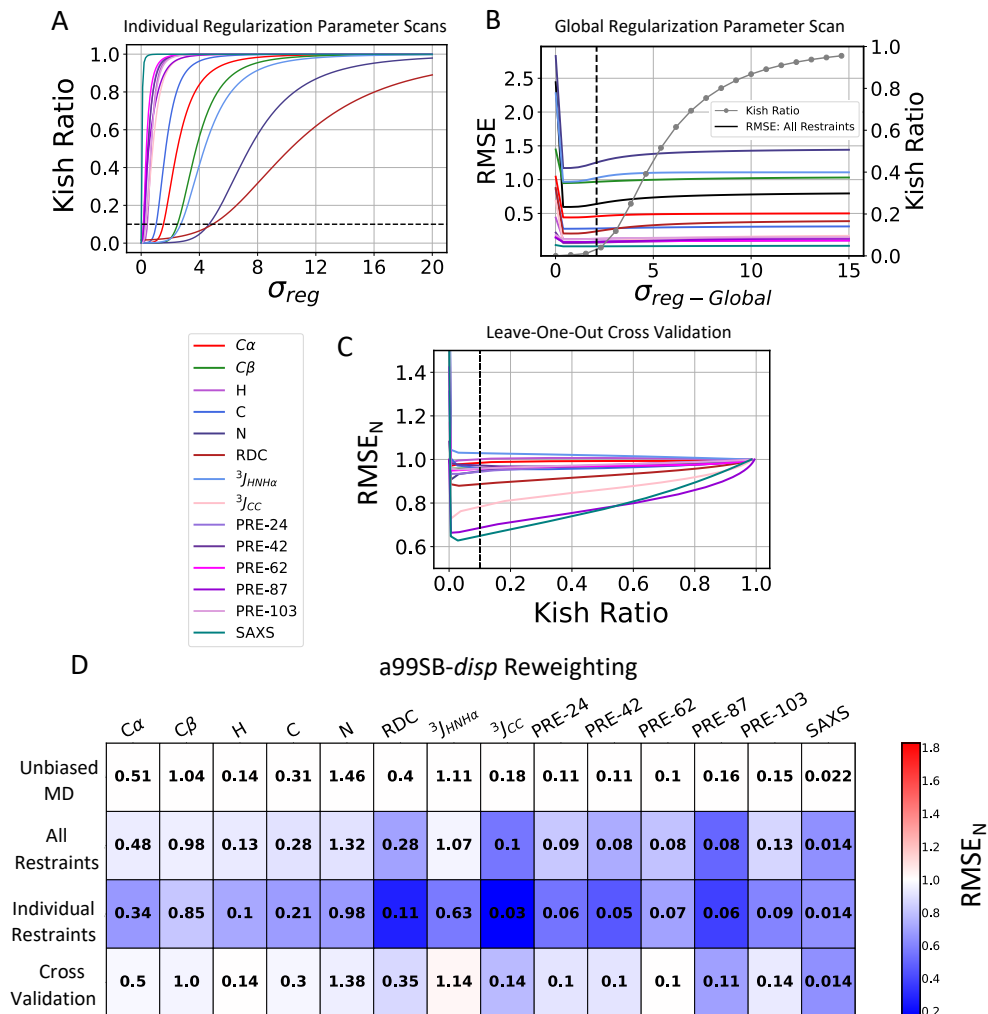
Supplementary Figure 17: Populations of β sheet conformations in unbiased and reweighted MD ensembles of PaaA2.



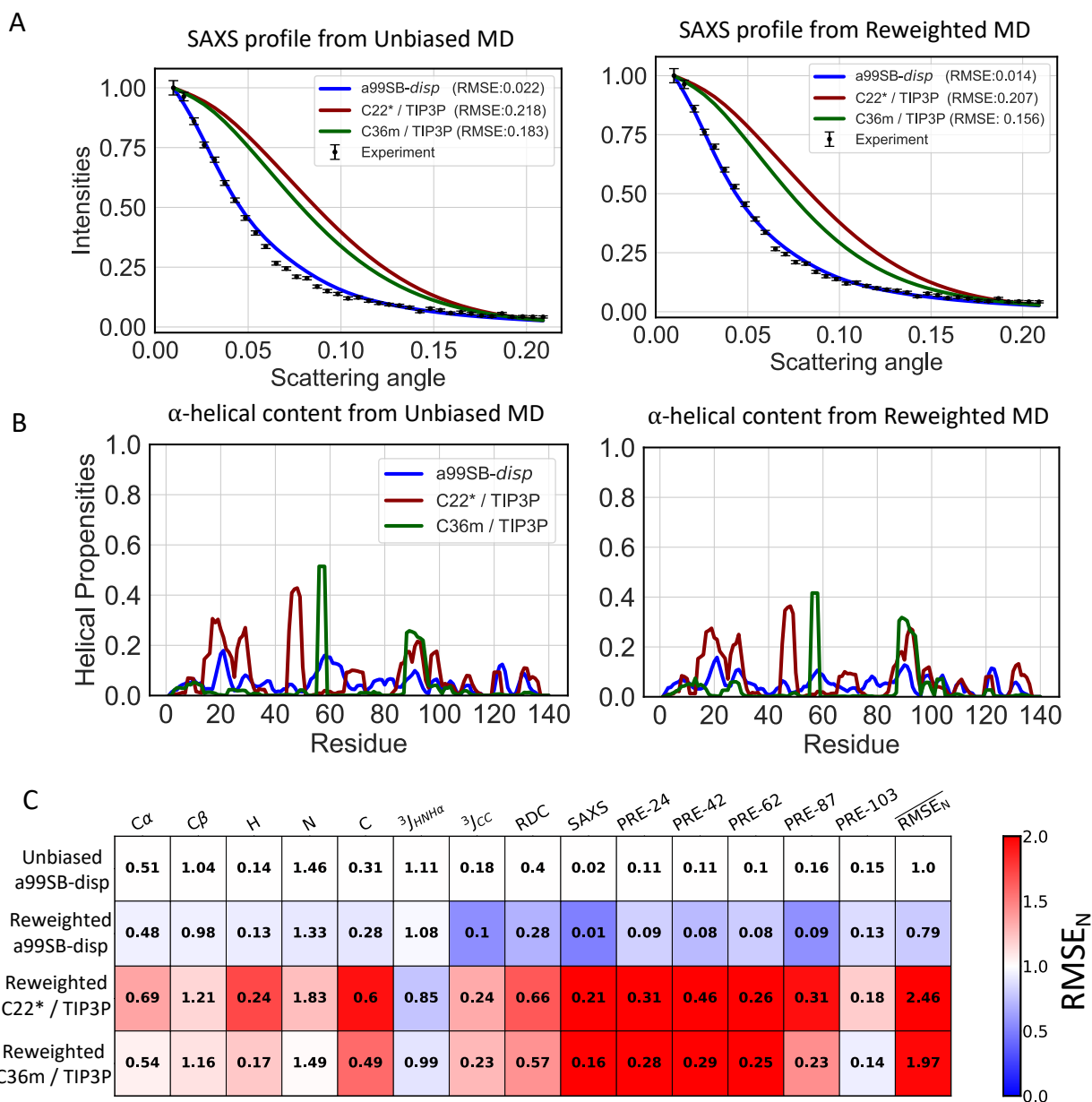
Supplementary Figure 18: Comparison of free energy surfaces of unbiased and reweighted MD ensembles of PaaA2. Free energy surfaces of unbiased MD ensembles (top row) and reweighted ensembles (bottom row) of PaaA2 are shown as a function of the radius of gyration (R_g) and α -helical order parameter $S\alpha$ of each conformation. The white dots represent the R_g and $S\alpha$ values of structures from a PaaA2 protein ensemble database (PED) ensemble calculated from experimental NMR and SAXS data.² Representative structures are shown for selected free energy basins of each ensemble. Structures are colored with a blue-to-red gradient from the N-terminus to the C-terminus.



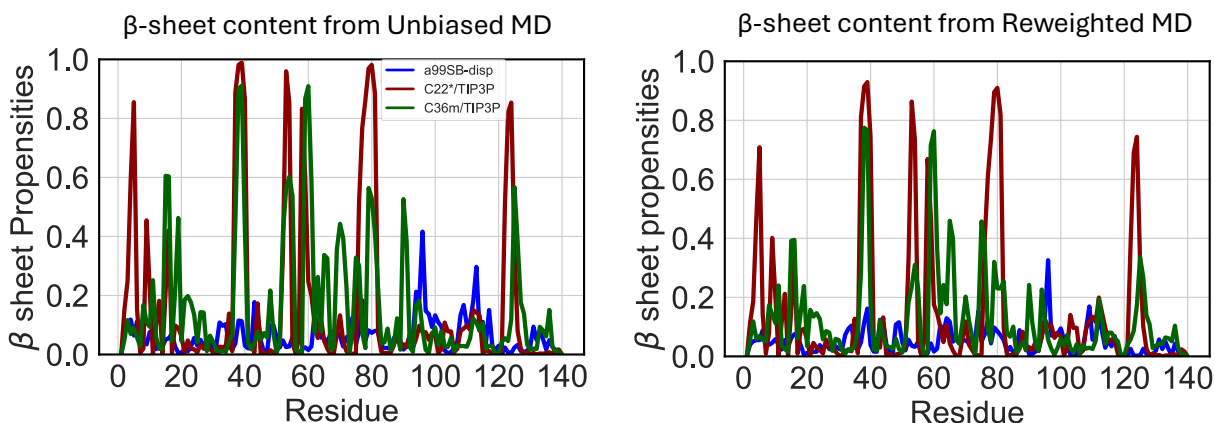
Supplementary Figure 19: Comparison of the populations of intramolecular contacts in unbiased and reweighted ensembles of PaaA2. Contact maps of unbiased MD ensembles (top row) and reweighted MD ensembles (bottom row) of PaaA2. Contacts between residues are defined using a 12 Å distance threshold between heavy atoms.



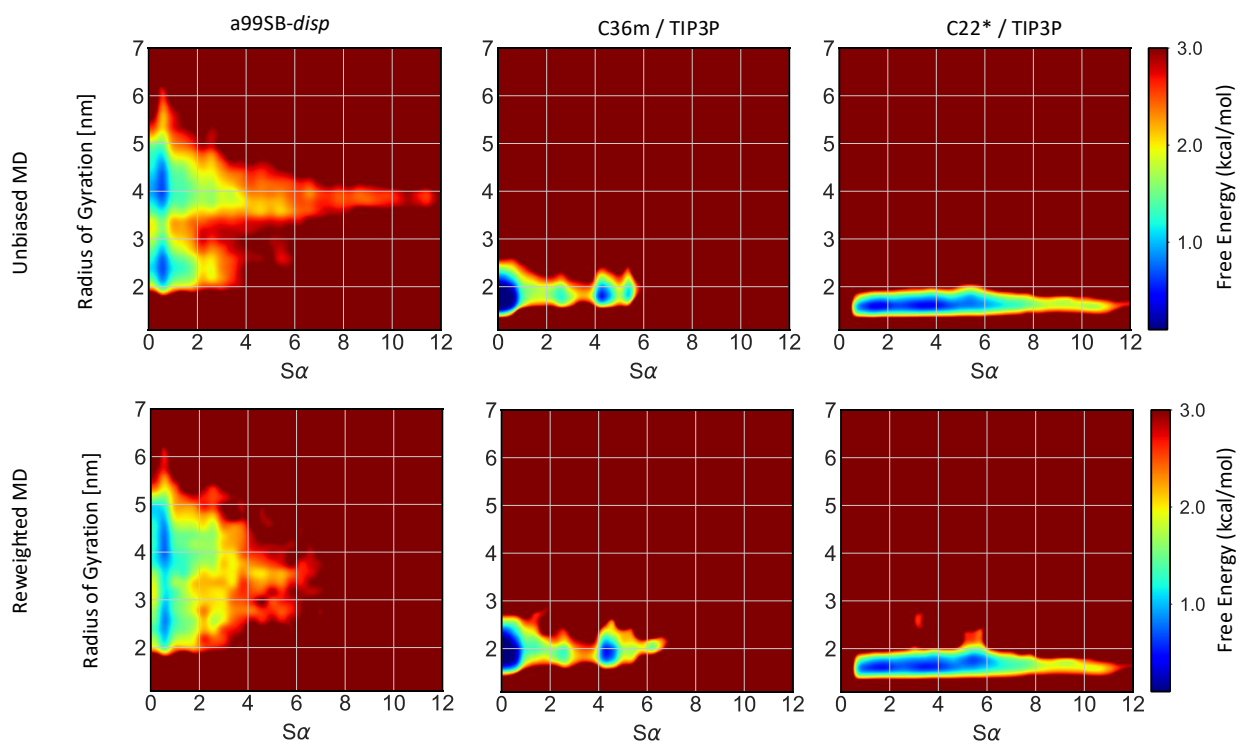
Supplementary Figure 20: Maximum entropy reweighting of an a99SB-*disp* MD ensemble of α -synuclein. **A)** The effective ensemble size of reweighted ensembles obtained using individual experimental data types as restraints as a function of the value of the regularization parameter σ_{reg} . **B)** Root mean square error (RMSE) between calculated and experimental data in reweighted ensembles as a function of the value of the global regularization parameter scaling factor ($\sigma_{reg-Global}$). The Kish ratio of each reweighted ensemble is indicated by gray dots. **C)** The normalized RMSE ($RMSE_N$) of unrestrained data in leave-one-out cross-validation tests where all other experimental data types are used as restraints for reweighting. $RMSE_N$ values of each data type are normalized by the RMSE observed in the unbiased a99SB-*disp* α -synuclein MD ensemble. Black dotted lines indicate the Kish ratio threshold $K = 0.10$ in each plot. **D)** RMSE of experimental data in reweighted ensembles obtained by reweighting the a99SB-*disp* MD ensemble with different combinations of restraints. RMSEs are displayed for the unbiased MD ensemble (“Unbiased MD”), a reweighted ensemble obtained using all experimental data as restraints (“All Restraints”), reweighted ensembles obtained using each individual data type as the only restraint (“Individual Restraints”) and reweighted ensembles obtained using all but one data type as restraints (“Cross Validation”). All reweighted ensembles were selected using a Kish ratio threshold of $K = 0.10$. Each square is colored to reflect the value of $RMSE_N$ relative to the unbiased a99SB-*disp* MD ensemble.



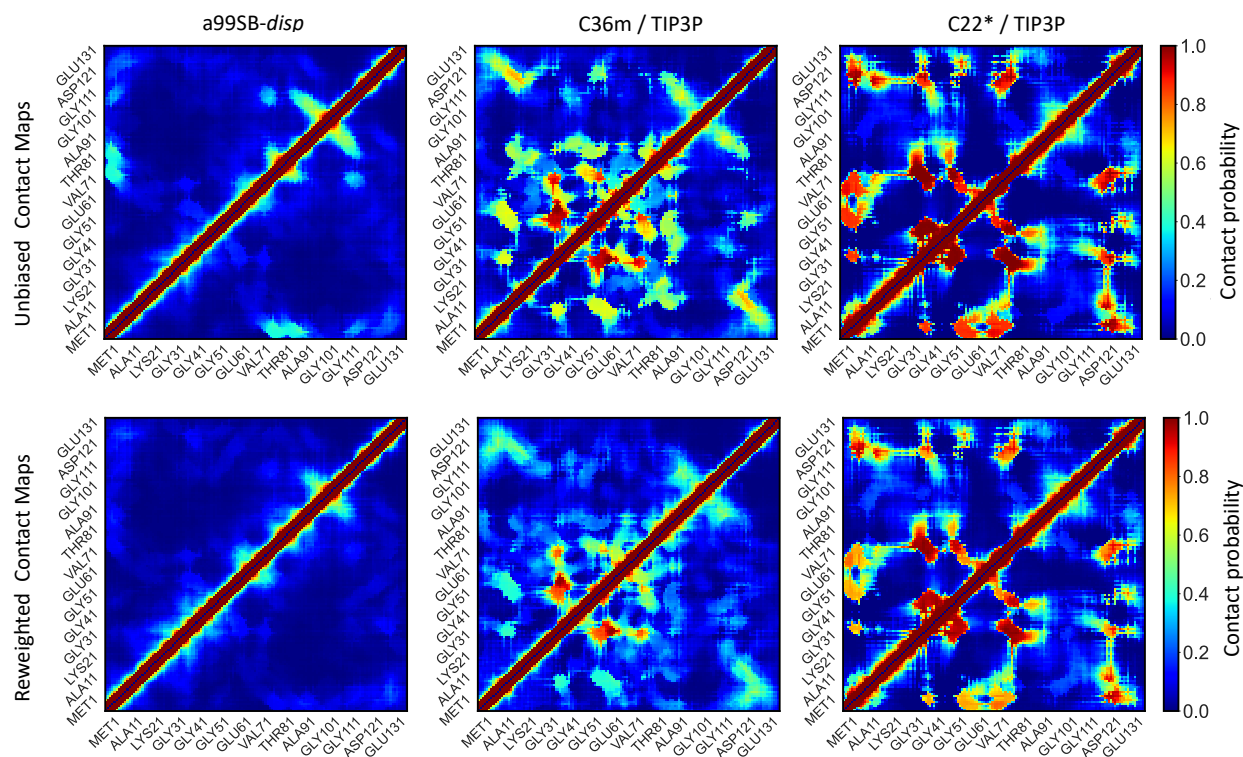
Supplementary Figure 21: Comparison of unbiased and reweighted MD ensembles of α -synuclein. **A)** Comparison of SAXS profiles calculated from unbiased MD ensembles and reweighted MD ensembles of α -synuclein. Experimental SAXS profile of α -synuclein is displayed with error bars as reported in the original publication³ **B)** Populations of α -helical conformations in unbiased MD ensembles and reweighted MD ensembles of α -synuclein. **C)** Comparison of the RMSE between calculated and experimental data in reweighted α -synuclein ensembles derived from different force fields. Reweighted ensembles were calculated using all experimental data as restraints with a Kish ratio threshold of $K=0.10$. Each square is colored to reflect the value of the normalized RMSE (RMSE_N) relative to the unbiased a99SB-disp MD ensemble.



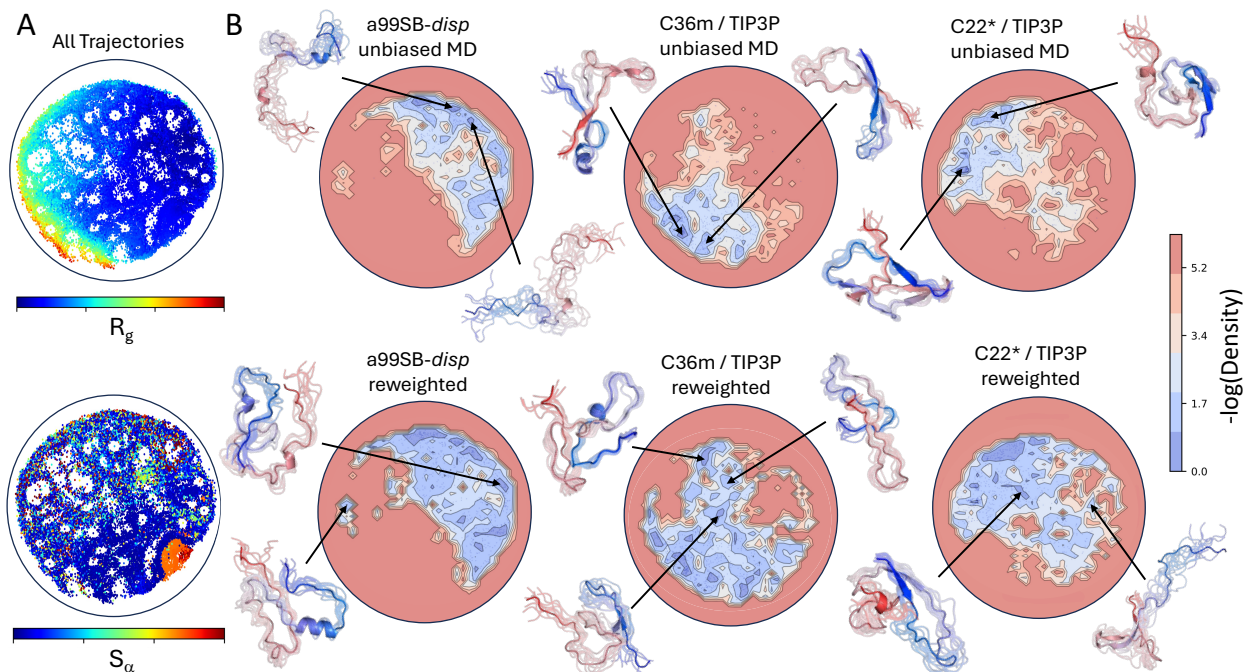
Supplementary Figure 22: Populations of β sheet conformations in unbiased and reweighted MD ensembles of α -synuclein



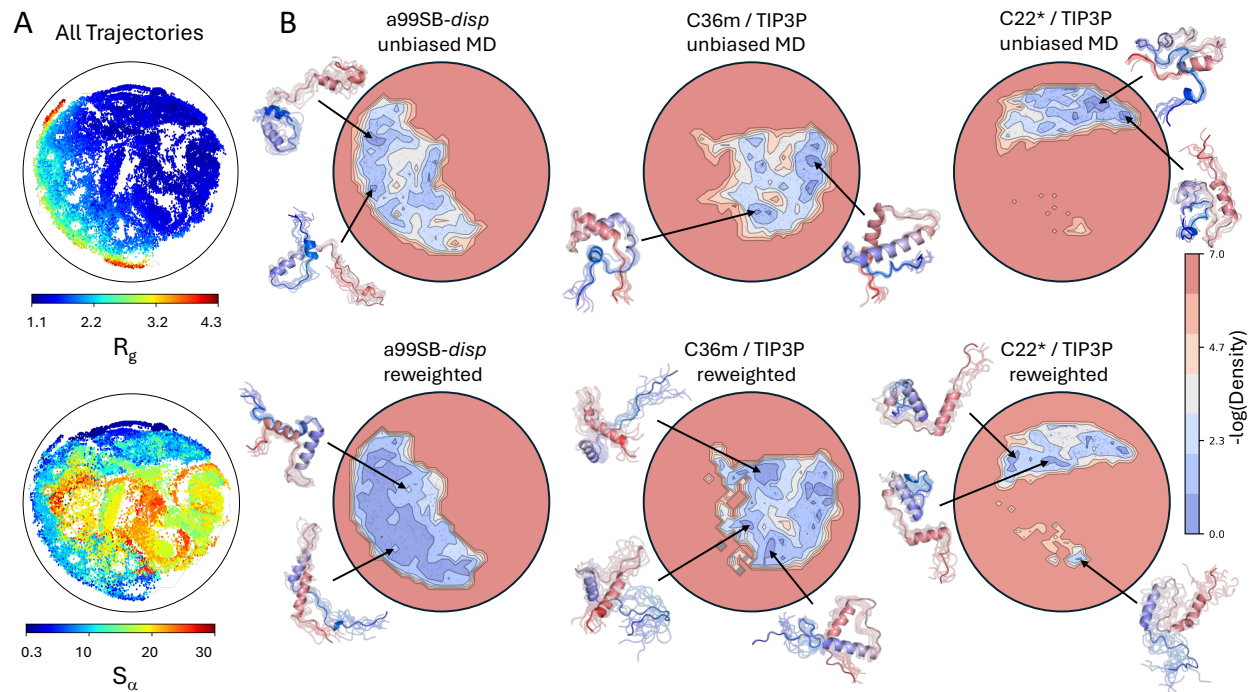
Supplementary Figure 23: Comparison of free energy surfaces of unbiased and reweighted ensembles of α -synuclein. Free energy surfaces of unbiased MD ensembles (top row) and reweighted ensembles (bottom row) of α -synuclein are shown as a function of the radius of gyration (R_g) and α -helical order parameter $S\alpha$ of each conformation.



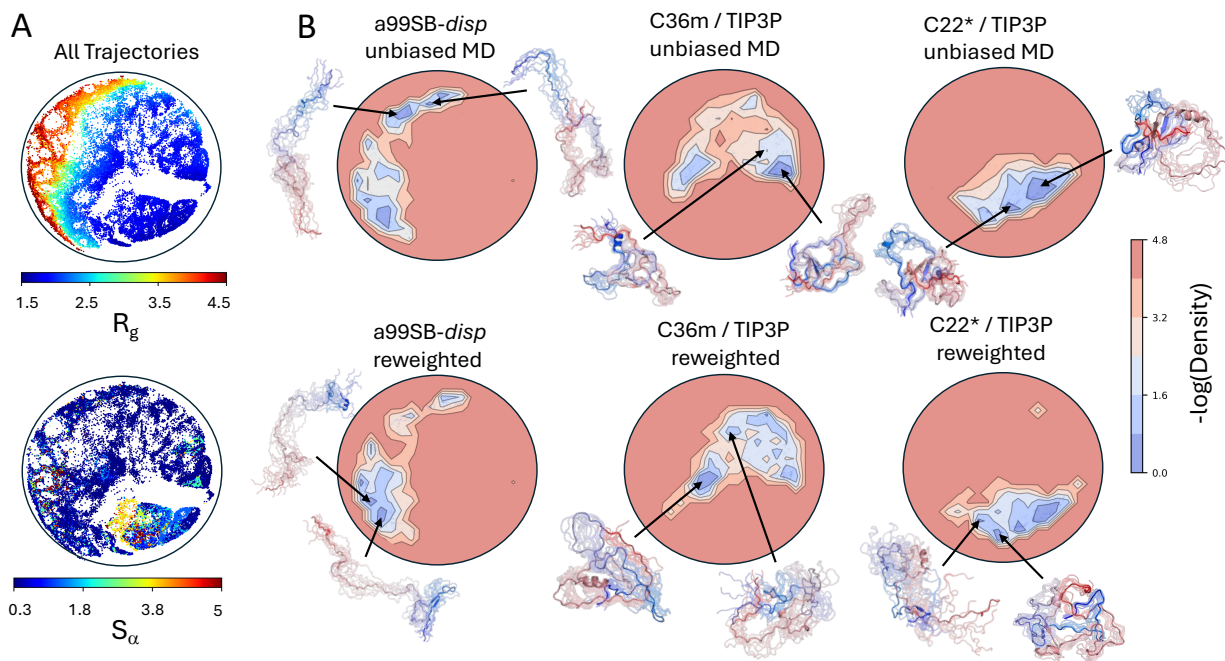
Supplementary Figure 24: Comparison of the populations of intramolecular contacts in unbiased and reweighted ensembles of α -synuclein. Contact maps of unbiased MD ensembles (top row) and reweighted MD ensembles (bottom row) of α -synuclein. Contacts between residues are defined using a 12 Å distance threshold between heavy atoms.



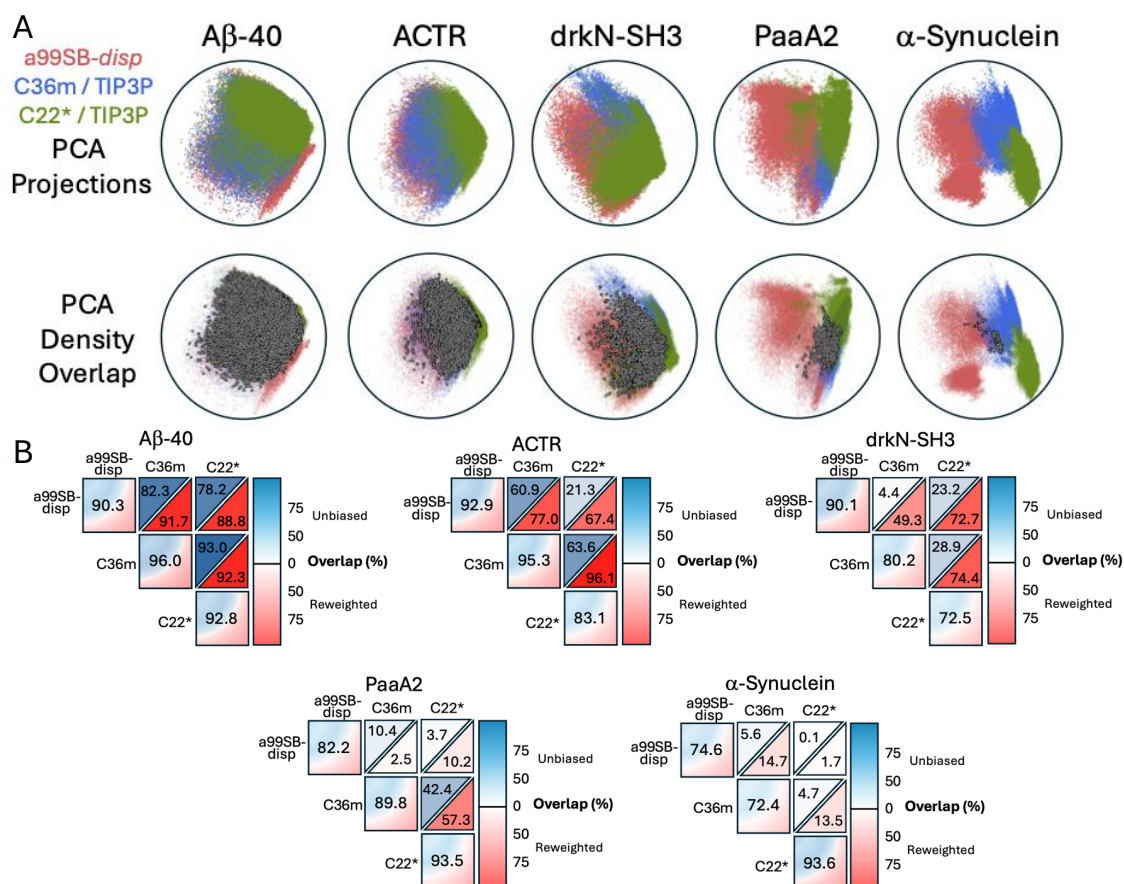
Supplementary Figure 25: Comparison of energy landscape visualization method (ELViM) projections of unbiased and reweighted MD ensembles of drkN SH3. A 2D ELViM embedding was calculated using all conformations from unbiased a99SB-*disp*, C22* and C36m MD ensembles of drkN SH3. **A)** The values of the radius of gyration (R_g) and the α -helical order parameter S_α of each conformation in the unbiased MD ensembles of drkN SH3 projected onto the ELViM latent space. **B)** Comparison of the density of ELViM projections of unbiased and reweighted MD ensembles of drkN SH3. Representative structures are shown for regions of high density in ELViM projections of each ensemble. Structures are colored with a blue-to-red gradient from the N-terminus to the C-terminus.



Supplementary Figure 26: Comparison of energy landscape visualization method (ELViM) projections of unbiased and reweighted MD ensembles of PaaA2. A 2D ELViM embedding was calculated using all conformations from unbiased a99SB-*disp*, C22* and C36m MD ensembles of PaaA2. **A)** The values of the radius of gyration (R_g) and the α -helical order parameter S_α of each conformation in the unbiased MD ensembles of PaaA2 projected onto the ELViM latent space. **B)** Comparison of the density of ELViM projections of unbiased and reweighted MD ensembles of PaaA2. Representative structures are shown for regions of high density in ELViM projections of each ensemble. Structures are colored with a blue-to-red gradient from the N-terminus to the C-terminus.



Supplementary Figure 27: Comparison of energy landscape visualization method (ELViM) projections of unbiased and reweighted MD ensembles of α -synuclein. A 2D ELViM embedding was calculated using all conformations from unbiased a99SB-*disp*, C22* and C36m MD ensembles of α -synuclein. **A)** The values of the radius of gyration (R_g) and the α -helical order parameter S_α of each conformation in the unbiased MD ensembles of α -synuclein projected onto the ELViM latent space. **B)** Comparison of the density of ELViM projections of unbiased and reweighted MD ensembles of α -synuclein. Representative structures are shown for regions of high density in ELViM projections of each ensemble. Structures are colored with a blue-to-red gradient from the N-terminus to the C-terminus.



Supplementary Figure 28: Comparison of the overlap of inter-residue distance PCA projections of unbiased and reweighted MD ensembles. A) 2D PCA projections were calculated for $A\beta$ 40, drkN SH3, ACTR, PaaA2 and α -synuclein. PCA projections were calculated for each protein using all conformations from unbiased a99SB-*disp*, C22* and C36m MD ensembles. Projections of unbiased a99SB-*disp*, C22* and C36m MD ensembles on the PCA latent space are shown in the top row. In the bottom row, each point is colored with an opacity proportional to its statistical weight after reweighting with all available experimental data. More transparent points corresponding to structures lower statistical weights. Histograms were estimated for each projection using 60 bins in each dimension. Data points belonging to histogram bins that contain at least one data point from each force field are colored dark gray. B) Overlap of PCA latent space kernel densities of unbiased and reweighted IDP ensembles. Values in the blue triangles reflect the PCA density overlap of unbiased MD ensembles derived from different force fields, values in red triangles reflect the PCA density overlap of reweighted ensembles derived from different force fields, and the diagonal elements reflect the PCA density overlap of the reweighted ensemble derived from each force field with its parent unbiased MD ensemble.

Supplementary Tables

Supplementary Table 1: Comparison of ELViM and PCA Silhouette Scores.

	α -Synuclein	PaaA2	ACTR	drkN-SH3	A β -40
ELViM	0.362	0.412	0.119	0.255	0.060
PCA	0.363	0.107	0.063	0.183	-0.034

Comparison of Silhouette scores⁴ (Eq. 13) to quantify the quality of low-dimensional projections of the conformational space sampled in MD simulations analyzed in this work.

Supplementary Discussion

Theory

In the following sections we provide a brief overview of the theory of the maximum entropy reweighting and the treatment of different sources error with regularization parameters. We focus on the formalism proposed by Bussi et al.⁵⁻⁷ which is the basis of our approach.

Overview of maximum entropy reweighting

A MD simulation generates a prior distribution $P_0(x)$ of conformations x . In initial implementations,^{6,8} maximum entropy reweighting approaches were designed to determine the distribution $P_1(x)$ which is the closest to $P_0(x)$ that fits *exactly* each experimental data point f_i^{exp} . To measure the statistical distance between the two distributions $P_0(x)$ and $P_1(x)$, one can use the Kullback–Leibler divergence, defined as:

$$D_{KL}[P_1|P_0] = \int P_1(x) \ln \frac{P_1(x)}{P_0(x)} dx \quad (1)$$

To determine $P_1(x)$, the closest distribution to $P_0(x)$ that fits the ensemble-averaged data,

the KL divergence is then minimized subject to the following two constraints:

$$\int P_1(x) f_i(x) dx - f_i^{exp} = 0 \text{ for } i = 0, 1, 2, \dots, M \quad (2)$$

$$\int P_1(x) dx = 1 \quad (3)$$

where $f_i(x)$ is the forward model needed to predict the value of observable i for each conformation x . In Eq. 2, each experimental observation constrains the ensemble average of the corresponding forward model $f_i(x)$ computed over the distribution $P_1(x)$ to be exactly equal to f_i^{exp} . The second constraint in Eq. 3 ensures that $P_1(x)$ is normalized.

The problem of finding the distribution $P_1(x)$ that is the best representation of the state of knowledge of a system after observing a set of experimental data can be cast in the framework of information theory. In this framework, amongst all the possible distributions compatible with the data, the one that has the highest degree of uncertainty should be selected in order to ensure minimal bias and that the data has been used as conservatively as possible. This uncertainty is expressed by the Shannon entropy⁹ of the random variable x with distribution $P_1(x)$ with respect to the reference distribution $P_0(x)$:

$$S[P_1|P_0] = - \int P_1(x) \ln \frac{P_1(x)}{P_0(x)} dx = -D_{KL}[P_1|P_0] \quad (4)$$

Maximizing the Shannon entropy in Eq. 4 to determine $P_1(x)$ is thus equivalent to minimizing the KL divergence in Eq. 1, hence the name of maximum entropy reweighting.

The solution for the minimization of the KL divergence subject to the constraints in Eq. 2 and 3 can be achieved by using Lagrange multipliers. We first define the following Lagrange function:

$$\mathcal{L} = D_{KL}[P_1|P_0] - \sum_{i=1}^M \lambda_i \left[\int P_1(x) f_i(x) dx - f_i^{exp} \right] - \mu \left[\int P_1(x) dx - 1 \right] \quad (5)$$

where λ_i and μ are Lagrange multipliers. There is one Lagrange multiplier λ_i for each of the

of the M experimental data points being used as constraints. To find stationary points of the Lagrangian, we set $\partial\mathcal{L}/\partial P_1 = 0$ and neglect the normalization factor. This leads to the following form for the desired distribution $P_1(x)$:

$$P_1(x) \propto e^{-\sum_{i=1}^M \lambda_i f_i(x)} P_0(x) \quad (6)$$

Approaches such as gradient descent or appropriate quasi-Newton methods can be utilized to calculate the values of the Lagrange multipliers, which ultimately yield to a new set of weights, one for each conformation of the ensemble:

$$w(x) \propto e^{-\sum_{i=1}^M \lambda_i f_i(x)} \quad (7)$$

Introducing different source of errors

By fitting exactly the experimental observations, the formalism introduced above neglects the fact that experimental data as well as forward models can be affected by random and systematic errors and therefore data should not be matched *exactly* to avoid overfitting to noise. In order to alleviate the problem of data overfitting, maximum entropy reweighting approaches can be modified to account for different sources of errors as well as a limited confidence in the prior distribution. Different approaches have been proposed over the years.^{6,10–13} In the formalism of Bussi et al.,⁶ errors are modeled by modifying the constraints of Eq. 2 by introducing an auxiliary variable for each data point ϵ_i , which represents the difference between experimental and predicted values. The new constraints are hence defined as follows:

$$\int P_1(x) (f_i(x) + \epsilon_i) dx - f_i^{exp} = 0 \text{ for } i = 0, 1, 2, \dots, M \quad (8)$$

Errors can be modeled by choosing a proper prior distribution function for the variables ϵ_i . We choose a Gaussian prior with fixed standard deviation σ_i for the i^{th} observable.

$$P_0(\epsilon_i) \propto \prod_{i=1}^M e^{-\left(\frac{\epsilon_i^2}{2\sigma_i^2}\right)} \quad (9)$$

The value of σ_i corresponds to the level of confidence in the i^{th} experimental data point and ultimately determines how well the reweighted ensemble will fit the data. $\sigma_i = \infty$ implies no confidence in the data and results in a reweighted ensemble that is identical to the unbiased simulation. $\sigma_i = 0$ implies complete confidence in the data and results in a reweighted ensembles in which the i^{th} data point is exactly matched, as in Eq. 2. The selection of appropriate values of the regularization parameters σ_i for each experimental data point is therefore essential to balance one's confidence in a prior model (ie. the accuracy of a force field), error estimates of experimental measurements and error estimates of the accuracy of forward models for predicting experimental data. Ultimately, the properties of the resulting conformational ensembles will be highly dependent on the relative magnitudes of the regularization parameters selected for each experimental data point.

Different strategies have been developed to identify the optimal values of σ_i . Typically, when performing reweighting using multiple types of experimental data, one value of σ_i is assigned to each class of experimental data, such as SAXS intensities or NMR chemical shifts, RDCs or scalar couplings. A grid search is then performed, usually independently for each class of experimental data, and the optimal value of σ_i is identified for each class to maximize the agreement of the reweighted ensemble with independent experimental data not used for reweighting.¹⁴ A final reweighting step using all the available experimental data is then performed.

Determining regularization parameters

In the approach proposed in this work, we assign one regularization parameter σ_i to each data point. To optimize the grid search over the space of σ_i , using a similar functional form to previous work¹⁵ we decompose this parameter into two contributions:

$$\sigma_i = \sqrt{\sigma_{i,MD}^2 + \sigma_{reg}^2}. \quad (10)$$

where σ_{reg} is the regularization parameter, one per data type, that describes both experimental and forward model errors of this specific type of experiment. The $\sigma_{i,MD}$ parameter, one per data point, represents both the statistical errors of forward models to predict experimental data and statistical errors that result from calculating average quantities over a finite-size ensemble. We calculate $\sigma_{i,MD}$ using the Flyvbjerg block analysis^{15,16} technique to account for the correlation between adjacent frames in the MD simulation or the standard error of the mean for uncorrelated data. We use a single free parameter, the (normalized) Kish effective sample size,^{7,17} or *Kish ratio* K , to determine the optimal values of σ_{reg} for each experimental data type as described in the main text. K is defined as

$$K = \frac{1}{N} \left(\sum_{i=1}^N w_i \right)^2 / \sum_{i=1}^N w_i^2 \quad (11)$$

where w_i is value of the statistical weight of the i th frame.

Developing a maximum entropy reweighting protocol with a single free parameter

Here, we summarize the key features of our approach and provide illustrative examples that highlight the rationale behind the choices made. We build upon the formalism of Bussi and coworkers,⁵⁻⁷ where the strength of the restraint on each experimental data point i is

determined by the value of an adjustable regularization parameter σ_i (Eq. 10). A key feature of our reweighting procedure is the use of effective ensemble sizes of reweighted ensembles to select the values of regularization parameters for reweighting. We quantify the effective ensemble size of reweighted ensemble using the Kish ratio^{7,17} (Eq. 11). Briefly, the Kish ratio (K) is a measure of the fraction of frames of an original ensemble with significant statistical weights after reweighting. A reweighted ensemble where $\sim 90\%$ of the frames have statistical weights close to zero will have a Kish ratio of $K \approx 0.10$. Conversely, if an unbiased MD ensemble contains 10,000 frames, a reweighted ensemble with a Kish ratio $K \approx 0.01$ would contain approximately 100 frames with statistical weights substantially larger than zero.

We initially attempted to reweight IDP ensembles using a cross-validation approach to determine the relative strengths of restraints for different experimental data types. We performed reweighting using a single value of the regularization parameter σ for all experimental data points of a given type (ie. $C\alpha$ chemical shifts, $C\beta$ chemical shifts, SAXS intensities) and iteratively decreased the value of σ while monitoring: i) the root mean squared error (RMSE) between calculated and experimental data used as restraints; ii) the RMSE between calculated and experimental cross-validating data not used as restraints; iii) the Kish ratio of the reweighted ensemble. We sought to identify an optimal value of σ for each experimental data type corresponding to the point where further reducing σ caused agreement with unrestrained cross-validating experimental data to deteriorate. We found, however, that many experimental data types were highly correlated and the RMSE of several cross-validating data sets continued to improve as the Kish ratio approached zero (SI Figure 1).

We tested several Bayesian and heuristic approaches to infer the optimal values of regularization parameters for each type of experimental data by monitoring the agreement of restrained and unrestrained cross-validation data using empirical RMSE thresholds and L-curve or “elbow method” analyses.^{10,14,18} In many instances, we observed that these approaches produced reweighted ensembles with extremely small effective ensemble sizes, where only hundreds of structures from an initial MD ensemble of nearly 30,000 structures maintained

appreciable statistical weights after reweighting ($K < 0.01$). When reweighting ensembles derived from different force fields, we observed that these approaches produced reweighted ensembles with dramatic differences in their effective ensemble size, depending on the initial agreement of unbiased MD ensembles with experimental data. This complicated global comparisons of the accuracy and structural properties of reweighted ensembles derived from different force fields.

As we desire reweighted ensemble sizes that are large enough to provide reasonable statistics for the conformational properties of IDPs, such as the overall topology of backbone conformations and the relative positions of sidechains, we opted to utilize the value of the Kish ratio of reweighted ensembles to determine regularization parameters for each data type. To account for the errors of forward models used to calculate experimental data and the effects of calculating average quantities over a finite-size ensemble, we introduced an additional regularization parameter $\sigma_{i,MD}$ for each experimental data point i used for reweighting (Eq. 10). Inspired by previous work,¹⁵ We calculated $\sigma_{i,MD}$ by applying the Flyvbjerg block analysis method¹⁶ to the forward model predictions obtained from the unbiased MD trajectory.

With this addition, we obtain one regularization parameter for each experimental data point $\sigma_i = \sqrt{\sigma_{i,MD}^2 + \sigma_{reg}^2}$, where σ_{reg} is the same for all data of a given data type (ie. $\sigma_{reg-C\alpha}$, $\sigma_{reg-C\beta}$, $\sigma_{reg-SAXS}$).

To determine an optimal minimum Kish ratio threshold for reweighting IDP ensembles, we examined how the conformational properties of reweighted ensembles changed as a function of σ_{reg} for each data type. We sought to identify a minimum Kish ratio threshold where the conformational properties of reweighted ensembles, such as secondary structure content, radius of gyration (R_g) and populations of intramolecular contacts, remained relatively consistent as σ_{reg} was decreased and further improvements in the agreement between calculated and experimental data were largely achieved by increasing the sparsity of the reweighted ensemble. For each data type, as the strength of experimental restraints was increased (by decreasing σ_{reg}) we generally observed an initial regime where agreement with experimental

data was achieved by changing the statistical weights of the structures belonging to the most populated regions of conformational space sampled in unbiased MD simulations. As Kish ratios approached smaller values (ie. $K < 0.10$), improved agreement with experimental data was often achieved by favoring a smaller number of structures from regions of conformational space with unfavorable free energies in unbiased MD simulations.

An illustrative example of this behavior is shown in SI Figure 2. Here, we use only two experimental backbone NMR chemical shifts of a single residue as restraints to reweight the unbiased a99SB-*disp* MD ensemble of PaaA2. We examine how reweighted distributions of the predicted values of these chemical shifts change as we increase the strength of these two restraints. We observe that as we decrease σ_{reg} from 16.0 to 2.0, we improve agreement with experimental data by increasing the statistical weights of conformations near the most populated free energy basins of the original unbiased MD ensemble. When we further increase the strength of experimental restraints by setting $\sigma_{reg} = 1.0$, the Kish ratio drops to $K = 0.003$, corresponding to an effective reweighted ensemble size of approximately one hundred structures from an initial ensemble of 29,976 structures. Notably, we observe that very few of the conformations that maintain non-negligible statistical weights in the reweighted ensemble are located within the free energy basins of the original unbiased MD ensemble (SI Figure 2). This dramatic reduction in the effective ensemble size and reshaping of the free energy surface results in an extremely marginal improvement in the agreement between the calculated and experimental values of these chemical shifts relative to the reweighted ensemble obtained with $\sigma_{reg} = 2.0$, which has a Kish ratio $K = 0.11$. This marginal improvement is likely within the noise of chemical shift predictor SPARTA+.¹⁹

We performed similar analyses for each experimental data type, monitoring how the free energy surfaces of reweighted ensembles and agreement with unrestrained cross-validating data varied with σ_{reg} . Based on these analyses, we elected to select regularization parameters using a Kish ratio threshold of $K = 0.10$. This means we iteratively decrease σ_{reg} for each individual data type until the resulting reweighted ensemble has a Kish ratio $K < 0.10$, and

select the smallest value of σ_{reg} where $K \geq 0.10$ as the final σ_{reg} value for that data type. After establishing an initial σ_{reg} for each individual data type, we perform a global reweighting using all experimental data as restraints simultaneously. In order to ensure a suitable ensemble size for the final reweighted ensemble, we rescaled the regularization parameter of each data type by a global scaling factor $\sigma_{reg-Global}$. We iteratively decrease the value of $\sigma_{reg-Global}$ and select the minimum value of $\sigma_{reg-Global}$ that produces a reweighted ensemble with Kish ratio $K \geq 0.10$ as the final value for global reweighting using all experimental data as restraints. The full reweighting protocol is illustrated for the a99SB-*disp* MD ensemble of A β 40 in SI Figure 3. We note that the choice of an optimal Kish ratio threshold is ultimately a subjective decision that will depend on the initial ensemble size of unbiased ensembles, the desired level of agreement with experimental data and the the desired degree of statistical sampling of properties of interest in reweighted ensembles. We found that varying the Kish ratio threshold between $K = 0.05$ and $K = 0.25$ did not appreciably change the structural properties of reweighted IDP ensembles or any of the conclusions of this study.

A maximum entropy reweighting protocol with a single free parameter improves agreement with extensive experimental datasets with minimal overfitting

We illustrate the key components of the proposed reweighting procedure by reweighting the unbiased a99SB-*disp* MD ensemble of A β 40 (SI Figure 3). We illustrate the process of determining σ_{reg} for each experimental data type of A β 40 in SI Figure 3A and the process of determining the global regularization parameter scaling factor $\sigma_{reg-Global}$ in Figure SI Figure 3B. In both plots, we denote the selected Kish ratio threshold $K = 0.10$ with a dashed line. The proposed automated reweighting procedure simultaneously improves agreement with all types of experimental data in this system without requiring any manual adjustments of the relative strength of experimental restraints. We observe that the RMSE between calculated

and experimental data monotonically decreases for all types of experimental data as the effective ensemble size of the reweighted A β 40 ensemble decreases from 30,000 to 3000, corresponding to the selected Kish Ratio threshold of $K=0.10$ (SI Figure 3A). We find that the RMSE between calculated and experimental data continues to monotonically decrease for all types of data until the effective ensemble size reaches ~ 150 structures ($K \sim 0.005$). This demonstrates that while we select a Kish Ratio threshold of $K=0.10$ to obtain a desired degree of statistical sampling in our final reweighted A β 40 ensemble, we could further improve global agreement with experimental data if we were willing to produce a sparser reweighted ensemble with less statistical sampling of conformational properties of interest. We note the RMSE between calculated and experimental data begins to deteriorate for extremely small effective ensemble sizes containing fewer than 70 structures, which we attribute to instability in the L-BFGS minimizer used to minimize the Lagrange function in SI Eq. 5 in the presence of 237 experimental restraints.

A key feature of the reweighting protocol proposed here is that the relative strengths of restraints for different types of experimental data are determined automatically. It is therefore important to determine if the proposed procedure simultaneously improves agreement with all data types in large experimental NMR and SAXS datasets or if the procedure is overfit to a subset of experimental data. To assess this, we performed i) reweighting using each experimental data type as the only restraint; ii) a global reweighting using all experimental data as restraints; and iii) a leave-one-out cross-validation reweighting using all but one experimental data type as restraints. The results of these reweighting calculations are shown for the a99SB-*disp* MD ensemble of A β 40 in SI Figure 3C and SI Figure 3D. In SI Figure 3D we compare the accuracy of reweighted ensembles to the unbiased ensemble using a normalized RMSE (RMSE_N) where we normalize the value of the RMSE calculated for each experimental data type by the corresponding value in the unbiased a99SB-*disp* MD ensemble ($\text{RMSE}_N = \text{RMSE}/\text{RMSE}_{\text{a99SB-disp}}$).

We observe that when we use all available experimental NMR data ($C\alpha$, $C\beta$, $H\alpha$, H , and N

backbone NMR chemical shifts, backbone $^3J_{HNH\alpha}$ coupling constants, and backbone RDCs) as restraints our reweighting protocol simultaneously improves agreement with all experimental data types (SI Figure 3D, “All Restraints”). For each experimental data type, we obtain a smaller RMSE between calculated and experimental data when we perform reweighting using only those data as restraints (SI Figure 3D “Individual Restraints”) compared to the RMSE obtained when performing reweighting using all available data as restraints. This demonstrates that satisfying multiple types of data requires trade offs in the agreement with each type of restraint and that our reweighting procedure naturally balances the strengths of different types of restraints without requiring manual adjustment of regularization parameters or cross-validation thresholds.

We observe that the RMSE between calculated and experimental data for each data type withheld in leave-one-out cross validation reweighting calculations is very similar to the RMSE obtained when reweighting using all experimental data as restraints (SI Figure 3D, “Cross validation”). This suggests that using a Kish ratio threshold of $K \geq 0.10$ to determine regularization parameters for each data type results in minimal overfitting when all experimental data are used as restraints in this system. We observe that the structural properties (secondary structure populations, intramolecular contact populations and R_g distribution) of the A β 40 conformational ensemble obtained by reweighting with all experimental data and the structural properties of conformational ensembles obtained from leave-one-out cross-validation reweighting calculations were largely indistinguishable (data not shown). This suggests that we are in a data rich regime where all of the reweighted ensembles of A β 40 derived from the unbiased a99SB-*disp* A β 40 MD ensemble are extremely similar.

We note that for the a99SB-*disp* A β 40 unbiased MD ensemble, with the exception of backbone RDCs, the agreement of unrestrained cross-validating data with experiment continues to improve as the effective ensemble size is decreased beneath 3000 structures, below the selected Kish Ratio threshold of $K=0.10$ (SI Figure 3C). We find that unrestrained

cross-validating continues to improve until the effective ensemble size of the reweighted ensemble contains fewer than 70 structures, demonstrating that we could further improve agreement with unrestrained cross-validating data if we were willing to obtain a much sparser final reweighted ensemble. This suggests that using a Kish ratio threshold of $K \geq 0.10$ to determine regularization parameters for each data type results in minimal overfitting when all experimental data are used as restraints.

Assessing the accuracy of IDP conformational ensembles

IDP conformational ensembles calculated using the maximum entropy reweighting procedure proposed in this work have exceptional agreement with extensive experimental NMR and SAXS datasets that describe the local and global structural properties of IDPs compared to the agreement observed in unbiased MD simulations performed with state-of-the-art force fields.^{20,21} We compare the accuracy of reweighted conformational ensembles of A β 40, drkN SH3, ACTR, PaaA2 and α -synuclein derived from 30 μ s unbiased a99SB-*disp*, C22* and C36m MD simulations in Figure 1, SI Figure 12, Figure 2, Figure 3 and SI Figure 21, respectively.

We observe similar agreement between calculated and experimental data in the reweighted a99SB-*disp*, C22*, and C36m ensembles of A β 40 (Figure 1), ACTR (Figure 2), and drkN SH3 (SI Figure 12). The $\overline{\text{RMSE}}_{\text{N}}$ of reweighted ensembles of A β 40, are 0.73, 0.63, 0.61, for a99SB-*disp*, C22* and C36m, respectively (Figure 1E); the $\overline{\text{RMSE}}_{\text{N}}$ of the reweighted ensembles of ACTR are 0.84, 0.84, 0.85, for a99SB-*disp*, C22* and C36m, respectively (Figure 2C); and the $\overline{\text{RMSE}}_{\text{N}}$ of reweighted ensembles of drkN SH3 are 0.80, 0.95, 0.91, for a99SB-*disp*, C22* and C36m, respectively (SI Figure 12C). All of the reweighted ensembles of these three IDPs are in exceptionally good agreement with experimental NMR and SAXS data, and are in substantially better agreement with experimental data than the most accurate unbiased MD ensembles of these proteins reported in previous benchmark studies.^{20,21}

We observe larger differences between calculated and experimental data in reweighted en-

sembles of PaaA2 (Figure 3) and α -synuclein (SI Figure 21). The $\overline{\text{RMSE}}_N$ of reweighted ensembles of PaaA2 are 0.80, 1.39, 1.14, for a99SB-*disp*, C22* and C36m, respectively (Figure 3C), while the $\overline{\text{RMSE}}_N$ of reweighted ensembles of α -synuclein are 0.79, 2.46, 1.97, for a99SB-*disp*, C22* and C36m, respectively (SI Figure 21C). Reweighted a99SB-*disp*, C36m and C22* ensembles of PaaA2 have a similar agreement with experimental NMR chemical shift and RDC data (Figure 3C). The agreement between calculated and experimental data of reweighted C22* and C36m α -synuclein ensembles is, however, substantially worse than the agreement observed in the reweighted a99SB-*disp* ensemble (SI Figure 21C).

The agreement with experimental SAXS data is particularly poor in reweighted C22* and C36m ensembles of PaaA2 and α -synuclein. The C22* and C36m MD ensembles of these IDPs are substantially more compact than the a99SB-*disp* ensembles both before and after reweighting (Figure 3, SI Figure 18, SI Figure 21, SI Figure 23). Reweighting C22* and C36m ensembles of PaaA2 and α -synuclein only produces a marginal improvement in their agreement with the experimental SAXS data (Figure 3A, SI Figure 21A). This results from the fact that unbiased C36m and C22* ensembles of PaaA2 and α -synuclein contain almost no conformations with a R_g as large as the experimental R_g determined from SAXS^{2,3,20} (SI Figure 18, SI Figure 23). This places an upper limit on the agreement that can be obtained from reweighting regardless of the specified Kish ratio threshold. Reweighted C36m and C22* ensembles of α -synuclein also have substantially worse agreement with experimental NMR chemical shift, RDC, and PRE data than the reweighted a99SB-*disp* ensemble, due to an overestimation of β -sheet content and overly compact ensemble dimensions (SI Figures 21-24).

We present detailed leave-one-out cross validation analyses of reweighting calculations performed on a99SB-*disp* MD ensembles of A β 40, drkN SH3, ACTR, PaaA2 and α -synuclein in SI Figure 3, SI Figure 11, SI Figure 7, SI Figure 16 and SI Figure 20, respectively. We observe similar global improvements in the agreement between calculated and experimental data in all proteins, illustrating the robustness of the reweighting protocol proposed here on

a range of IDPs. We note one exception in the case of the protein ACTR, where the agreement between calculated and experimental SAXS data obtained using all experimental data as restraints is slightly (6%) worse than the agreement observed in the unbiased a99SB-*disp* MD ensemble of ACTR (SI Figure 7). We also observe that the RMSE between calculated and experimental SAXS data is 18% worse than the RMSE of the unbiased a99SB-*disp* MD ensemble when SAXS data is excluded as cross-validating data. This demonstrates some incompatibility in the agreement between experimental NMR data (backbone chemical shifts, backbone RDCs and PREs) and SAXS data in the conformations sampled in the unbiased a99SB-*disp* MD ensemble of ACTR.

Remarkably, only three of the fifty experimental data types withheld in leave-one-out cross-validation reweighting calculations performed on a99SB-*disp* MD ensembles of A β 40, drkN SH3, ACTR, PaaA2 and α -synuclein were found to have worse agreement with experiment in reweighted ensembles than in unbiased MD ensembles: PRE data measured with a spin-label on residue 59 of drkN SH3 (drkN SH3 PRE-59), ACTR SAXS data and α -synuclein backbone $^3J_{HNH\alpha}$ coupling constants (SI Figure 11, SI Figure 7, SI Fig 20). Similarly, only five data types withheld in leave-one-out cross-validation reweighting tests performed on C36m ensembles had worse agreement with experiment in reweighted ensembles than in unbiased MD ensembles (ACTR PRE-3, ACTR PRE-41, ACTR PRE-61, ACTR SAXS data and α -synuclein PRE-103) and only three data types withheld in leave-one-out cross-validation reweighting tests performed on C22* ensembles had worse agreement in reweighted ensembles (drkN SH3 PRE-59, ACTR PRE-41 and α -synuclein PRE-103). Together, these results demonstrate that the automated maximum entropy reweighting protocol introduced in this investigation to refine IDP ensembles improves agreement with extensive experimental datasets from NMR spectroscopy and SAXS with minimal overfitting.

Comparing IDP conformational ensembles derived from different force fields

The populations of α -helices in unbiased and reweighted ensembles of ACTR and PaaA2 are shown in Figure 2 and Figure 3, respectively. The populations of α -helices in unbiased and reweighted ensembles of A β 40, drkN SH3 and α -synuclein are shown in SI Figure 4, SI Figure 12, and SI Figure 21, respectively. We observe highly consistent populations of moderately stable helices (populations of 20%-40%) in the reweighted ensembles of ACTR and drkN (Figure 2, SI Figure 12). In the case of PaaA2, we observe large differences in the locations and populations of helical elements in unbiased MD ensembles that become substantially more similar after reweighting (Figure 3). We note that there are strong signals for the helical elements of drkN SH3, ACTR and PaaA2 in experimental NMR chemical shift, RDC, and NOE data and that the populations of helical conformations observed in our reweighted ensembles are in excellent agreement with helical populations directly estimated from experimental data.^{2,22-24}

The populations of β -sheets in unbiased and reweighted ensembles for A β 40, drkN SH3, ACTR, PaaA2 and α -synuclein are shown in Figure 1, SI Figure 13, SI Figure 8, SI Figure 17 and SI Figure 22 respectively. We observe that all reweighted ensembles of A β 40 converge to very similar secondary structure populations, with residues 15-20 and residues 30-35 forming β -sheets with populations of 20%-40% (Figure 1). We observe several highly populated β -sheets (populations of 40%-95%) in unbiased C36m and C22* MD ensembles of drkN SH3 and α -synuclein that produce notable discrepancies with experimental NMR data (SI Figure 12, SI Figure 13, SI Figure 21, SI Figure 22). The populations of these β -sheets are only partially decreased upon reweighting. The unbiased C36m MD ensemble of drkN SH3 contains β -sheets between by residues 1-10 and residues 20-30 with populations of 40%-80%. These populations are reduced to 20%-40% after reweighting with all experimental data. The reweighted C36m drkN SH3 ensemble, however, has substantially worse agreement with NMR chemical shifts and $^3J_{HNH\alpha}$ scalar couplings than the reweighted a99SB-*disp* drkN SH3

ensemble, which contains very little residual β -sheet structure. We observe similarly modest reductions in spurious populations of β -sheets in reweighted C22* and C36m ensembles of α -synuclein (SI Figure 21, SI Figure 22).

To determine if the persistence of spurious β -sheet populations is the result of our selected Kish ratio threshold of $K = 0.10$ we computed reweighted C22* and C36m ensembles of drkN SH3 and α -synuclein using a Kish ratio threshold of $K = 0.01$. We observed similar β -sheet populations in these reweighted ensembles, suggesting that the persistence of these β -sheets is likely the result of the relatively weak effect of residual β -sheet structure on experimental secondary NMR chemical shifts^{25,26} and the properties of empirical structure based chemical predictors. Previous work has shown that experimental secondary NMR chemical shifts and structure-based backbone NMR chemical shift predictions from algorithms such as SHIFTX+²⁷ and SPARTA+¹⁹ are inherently less discriminative of conformations in the β -sheet region of backbone ϕ/ψ Ramachandran space than conformations in the α -helical region.^{25,26}

We find that our proposed reweighting procedure only marginally reduces the populations of potentially spurious β -sheets in reweighted ensembles of drkN SH3 and α -synuclein relative to those calculated by X-EISD²⁸ and ASTEROIDS.³ We note that these approaches, and other ensemble calculation approaches that generate conformations by statistically sampling backbone dihedral angles, may not sample any hydrogen-bonded β -sheet conformations in initial conformational pools used for ensemble selection, but rather aim to identify elevated sampling of β -sheet Ramachandran space relative to statistical coil models. Therefore, populations of β -sheets observed in reweighted MD ensembles may not be directly comparable to β -propensities identified in these ensemble calculation approaches. Nevertheless, we caution that the presence of substantially populated β -sheets in reweighted MD ensembles produced by our reweighting procedure does not necessarily suggest strong experimental evidence for residual β -sheet populations. We encourage a direct inspection of agreement with experimental data for these β -sheet elements. We emphasize that an important difference between

ensembles obtained from reweighting long timescale MD simulations and ensemble selection approaches such as ASTEROIDS^{3,25} or EOM²⁹ is that the positions of side chains in reweighted MD ensembles are governed by the physics of the underlying force fields, compared to a stochastic sampling of side chain rotamer libraries to place side chains and avoid steric clashes. The positions of sidechains in reweighted MD ensembles may therefore have more physical meaning.

We observe that IDP ensembles derived from MD simulations run with different force fields converge to very similar populations of residual helical structure after reweighting, but IDP ensembles with identical populations of α -helices can have large differences in the lengths of helical elements observed in individual conformations. An IDP ensemble containing ten consecutive residues with 50% α -helical populations might be constituted of only structures containing a contiguous ten-residue helix or structures with no helical content. Alternatively, this ensemble could contain conformations with short helical elements where the populations of helical turns are relatively uncorrelated from neighboring residues. To obtain a more detailed measure of the similarity of the individual conformations within unbiased and reweighted IDP ensembles, we compared the free energy surfaces of the unbiased and reweighted ensembles of each protein as a function of R_g and the α -helical order parameter $S\alpha$. $S\alpha$ is a measure of the number of seven residue fragments in a structure that resemble an ideal helix³⁰ (See Methods). The free energy surfaces of unbiased and reweighted ensembles of PaaA2 are shown as a function of R_g and $S\alpha$ in SI Figure 18. The free energy surfaces of unbiased and reweighted ensembles of A β 40, drkN SH3, ACTR and α -synuclein are shown as a function of R_g and $S\alpha$ in SI Figure 5, SI Figure 14, SI Figure 9 and SI Figure 23, respectively.

We observe dramatic differences in the free energy surfaces of the unbiased a99SB-*disp*, C36m and C22* ensembles of PaaA2 both before and after reweighting (SI Figure 18). The reweighted a99SB-*disp*, C22* and C36m PaaA2 ensembles have similar per-residue populations of helical conformations but have very little overlap when projected using R_g and $S\alpha$

descriptors. For comparison, we plot the R_g and $S\alpha$ values of 50 conformations contained in a protein ensemble database (PED)³¹ ensemble of PaaA2 calculated directly from NMR and SAXS restraints (PED00013)² as white dots in each free energy surface in SI Figure 18. We observe that the unbiased a99SB-*disp* PaaA2 MD ensemble has two broad free energy minima in the R_g and $S\alpha$ projection. One of these minima overlaps very well with the conformations in PED00013 ensemble, and the depth of this minimum is substantially increased upon reweighting. The overly compact unbiased C22* and C36m MD ensembles of PaaA2 have relatively little overlap with the experimental PaaA2 ensemble in this projection. Upon reweighting, the overlap with more compact conformations in the experimental ensemble increases, but the reweighted C22* and C36m ensembles are substantially more similar to the unbiased MD ensembles from which they were derived than to the experimentally restrained PED ensemble or the reweighted a99SB-*disp* PaaA2 ensemble. Reweighted PaaA2 ensembles clearly do not converge to similar conformational distributions across different force fields. Free energy surfaces of unbiased and reweighted drkN SH3 ensembles are compared to an experimental ensemble (PED00427) in SI Figure 14. All three unbiased MD ensembles of drkN SH3 have substantial overlap with the experimental ensemble in the R_g and $S\alpha$ projection, and the free energy surfaces of reweighted ensembles are relatively similar. In contrast to PaaA2, reweighted drkN SH3 ensembles appear to converge to similar conformational distributions with similar lengths of helical elements and a similar coupling between helix formation and chain compaction. Free energy surfaces of reweighted and unbiased ensembles of A β 40, ACTR and α -synuclein are shown as a function of R_g and $S\alpha$ in SI Figure 5, SI Figure 9 and SI Figure 23, respectively. Reweighted ensembles of A β 40, which have minimal helical content, are fairly similar in this projection. Unbiased ensembles of ACTR have fairly large differences in this projection, and become more similar after reweighting. Unbiased and reweighted α -synuclein ensembles have very little overlap in these projections.

We examine the similarity of the populations of intramolecular contacts, or *contact maps*, of unbiased and reweighted ensembles of A β 40, drkN SH3, ACTR, PaaA2 and α -synuclein

in SI Figure 6, SI Figure 15, SI Figure 10, SI Figure 19 and SI Figure 24, respectively. Unbiased A β 40 ensembles have a similar pattern of intramolecular contacts, defined by contacts between residues 15-20 and residues 30-35, with notable differences in populations. The populations of these contacts, which are driven by transient β -sheet formation, become more similar after reweighting (SI Figure 6). The populations of intramolecular contacts of drkN SH3 and ACTR are quite similar in unbiased MD ensembles of C36m and C22*, while the more extended unbiased a99SB-*disp* MD ensembles contain fewer intramolecular contacts. Upon reweighting ensembles of drkN SH3 and ACTR with experimental datasets that contain PREs and SAXS data, the contact maps become more similar, and the contact maps of the reweighted ACTR ensembles are particularly similar (SI Figure 10).

The reweighted C36m and C22* ensembles of drkN SH3 contain several intramolecular contacts resulting from β -sheet structure not observed in the reweighted a99SB-*disp* ensemble (SI Figure 15). The overly compact unbiased C36m and C22* MD ensembles of PaaA2 and α -synuclein contain a large number of highly populated intramolecular contacts relative to the more realistically expanded unbiased a99SB-*disp* MD ensembles (SI Figure 19, SI Figure 24). The contacts observed in C36m and C22* PaaA2 and α -synuclein ensembles remain highly populated after reweighting, as there are too few extended conformations in the unbiased MD ensembles to reduce the populations of these contacts.

Importance of accuracy and adequate sampling in initial ensembles used for reweighting

We emphasize that quality of reweighted ensembles obtained with the maximum entropy reweighting protocol proposed here, or obtained with any maximum entropy reweighting method, is highly dependent on the quality of the initial unbiased ensembles used for reweighting. As maximum entropy methods identify the minimal perturbation to an initial ensemble required to satisfy a given set of experimental restraints, the accuracy of reweighted ensembles is fundamentally limited by the compatibility of the structures in the initial en-

semble with experimental data. For example, if an initial ensemble of an IDP is generated by a short MD simulation with an inaccurate force field that only samples overly collapsed states with a substantially smaller radius of gyration than the experimental radius of gyration, no reweighting method can produce an ensemble in agreement with the experimental SAXS data. Similarly, if experimental NMR chemical shifts indicate the presence of highly populated helical elements in an IDP ensemble, and an initial ensemble used for reweighting contains no helical conformations in these regions, it will not be possible to produce a reweighted ensemble with the correct populations of helical elements, regardless of strategies used to determine regularization parameters.

The success of any reweighting protocol is therefore contingent on using reasonably accurate, well-sampled initial ensembles as input for reweighting. IDP conformational ensembles are highly heterogeneous and contain structures with many distinct topologies separated by large free-energy barriers. Sampling this vast conformational space in explicit solvent all-atom MD simulations is extremely challenging. Therefore, even if an accurate force field and water model combination is used for MD simulations^{20,21,32-34} very long-timescale MD simulations (10s-100s of μ s) or enhanced sampling methods³⁵⁻³⁸ are generally required to obtain statistically meaningful descriptions of IDP conformational ensembles.

We note that the unbiased ensembles of PaaA2 and α -synuclein obtained from the C22* and C36m force fields sample almost no conformations with an R_g as large as the experimental R_g determined from SAXS,^{2,3,20} due to the propensity of these force fields, used in combination with the TIP3P water model, to substantially underestimate the R_g of IDPs longer than 60 residues.²⁰ As a result, the reweighted ensembles of these proteins do not converge to similar conformational distributions, reflecting the inherent limitations of maximum-entropy reweighting methods discussed above.

Comparisons to previous work

The reweighting approach presented here builds upon several successful maximum entropy and Bayesian approaches for reweighting conformational ensembles of flexible molecules with experimental data.^{5-7,10-12,14} Refining conformational ensembles of flexible molecules with maximum entropy reweighting poses two inherent challenges: i) determining how to balance the strengths of experimental restraints against one’s confidence in the initial conformational ensemble, and ii) determining the relative strengths of different types of experimental restraints. Several heuristic and theoretical approaches have been proposed for determining or inferring the strengths of experimental restraints when reweighting ensembles of flexible molecules.^{10-14,18,28} When considering experimental datasets with a small number of experimental data types cross-validation or L-curve/elbow methods provide a practical and direct approach. When performing reweighting using a few sets of experimental data it is reasonable and expedient to determine the strength of each set of restraints by identifying the strength where the agreement of calculated cross-validating data with experiment begin to deteriorate.

When reweighting conformational ensembles with data from many different experiments this process becomes less straightforward and one must attempt to normalize the relative RMSE between predicted and experimental values based on estimates of experimental measurement errors and the errors of forward models used to predict experimental data. This is frequently done through the use of χ^2 values obtained by normalizing RMSE calculations with various error models. While it is possible to propose reasonable estimates of experimental measurement and forward model prediction errors for different types of data, these quantities are inherently uncertain and these estimates introduce a degree of subjectivity into reweighting protocols. Additionally, in cases where many experimental data are highly correlated, such as NMR data that report on the distribution of backbone dihedral angles of IDPs (ie. chemical shifts, scalar coupling constants, RDCs), cross-validation may not be a practical approach to determine the relative strengths of restraints.

We have proposed a heuristic approach for determining the relative strengths of experimental restraints that directly builds upon previous work that considers the Kish ratios of reweighted ensembles.⁷ This approach is motivated by several practical considerations. One motivation is a desire to objectively compare the accuracy of IDP ensembles derived from unbiased MD simulations run with force fields with different levels of accuracy. If a reweighting approach for calculating IDP ensembles contains adjustable parameters to determine the relative importance of different types of experimental data in a reweighting objective function, it is possible to adjust the relative “scores” or likelihoods of reweighted ensembles derived from different force fields by adjusting these parameters.

A second motivation for using the effective ensemble size of reweighted ensembles to choose the relative strengths of experimental restraints is a desire to directly compare the structural properties and similarity of reweighted ensembles derived from different force fields. If we choose to use RMSE (or χ^2) thresholds to determine the strengths of experimental restraints, we observe that reweighted ensembles derived from different force fields can achieve similar RMSEs with experimental data with dramatically different effective ensemble sizes. Requiring an overly compact C22* or C36m ensemble to agree with SAXS data within a specified tolerance could require reducing the ensemble to only tens of structures, while a more realistically extended a99SB-*disp* ensemble might contain 1,000s or 10,000s of structures after reweighting. Comparing the structural properties and structural overlap of reweighted ensembles with orders-of-magnitude differences in effective ensemble sizes is less meaningful than comparing ensembles with similar effective ensemble sizes.

An additional motivation for using the effective ensemble size of reweighted ensembles to determine the relative strengths of restraints is our desire to adequately represent the most populated regions of conformational space sampled in unbiased MD simulations in reweighted IDP ensembles. One application of our research efforts to determine accurate atomic resolution ensembles of IDPs is to obtain mechanistic insight into their binding mechanisms with small molecule drugs.^{39–44} These dynamic and heterogeneous binding mechanisms are medi-

ated by statistical distributions of backbone conformations and positions of sidechain pharmacophores. As such, IDP ensembles with a small number of conformations that sparsely sample backbone topologies and the relative positions of sidechain atoms are unlikely to be informative for predicting small molecule binding sites or understanding subtle shifts in the conformational ensembles of IDPs that occur upon small molecule binding.^{39,40,42–44}

We note a conceptual similarity of the reweighting procedure proposed here with the concept of gentle ensemble refinement in the recently published work of Köfinger and Hummer.¹³ In this work, the authors propose an elegant approach to balance the strength of experimental restraints in maximum entropy reweighting against confidence in an initial prior model. They relate the KL divergence between the weights of an initial and reweighted ensemble to an energy uncertainty of the unbiased ensemble, and propose to select the optimal value of restraint strength by relating the expected force field accuracy in the space of experimental observables to the expected energy variance upon reweighting. While motivated by theoretical considerations, this approach is similar in spirit and in practice to the heuristic approach proposed here, where we specify an acceptable limit to changes in the statistical weights of conformations in unbiased ensembles *a priori*. We also note that one could implement a similar heuristic approach to one proposed here in the context of the L-curve/elbow method. We draw attention to the small effective ensemble sizes obtained with our proposed reweighting procedure if we utilize agreement with restrained or unrestrained data as the primary criterion to select the strength of the experimental restraints. We observe that the RMSE between calculated and experimental data of both restrained and unrestrained cross-validating data largely improve as the strength of experimental restraints is increased until ensemble sizes become very small (Kish ratio $K < 0.01$), corresponding to effective ensemble sizes of a few hundred structures. We note that several approaches to calculate IDP conformational ensembles with a minimal set of structures, sometimes referred to as *maximum parsimony*⁴⁵ approaches, have previously found that ensembles of ~ 50 -500 structures are required to best reproduce NMR and SAXS data of IDPs.^{2,3,23,25,28,46} It appears that the IDP

ensembles obtained by reweighting long timescale MD simulations using cross-validation approaches produce ensembles with a similar size to ensembles calculated by minimal ensemble approaches.

We observe that the reweighted ensemble of IDPs obtained in this investigation have similar structural properties to conformational ensembles calculated from NMR and SAXS data using minimal ensemble or maximum parsimony approaches. We find that α -helical populations of reweighted drkN SH3 ensembles obtained in this work agree well with the helical populations of drkN SH3 ensembles derived from NMR and SAXS data with approaches such as ENSEMBLE⁴⁶ and X-EISD.²⁸ We also observe that α -helical populations in PaaA2 ensembles obtained from our reweighting approach agree reasonably well with populations obtained in a previously reported PaaA2 ensemble² derived from NMR and SAXS data using the EOM approach.²⁹

Supplementary Methods

Experimental data used for reweighting

We used previously reported experimental datasets containing different combinations of experimental NMR and SAXS data to reweight simulations of each protein. Backbone NMR chemical shifts^{2,22,23,47,48} and residual dipolar couplings (RDCs)^{2,23,24,49,50} were used as restraints for all proteins. Three-bond backbone 3J scalar coupling constants were used as restraints for A β 40⁵¹ and α -synuclein⁵² ensembles. Paramagnetic relaxation enhancements (PREs) were used as restraints for drkN SH3,⁴⁶ ACTR,²³ and α -synuclein^{49,53,54} ensembles. SAXS data were used as restraints for drkN SH3,⁵⁵ ACTR,⁵⁶ PaaA2,² and α -synuclein³ ensembles. We treat the NMR chemical shifts of each backbone nucleus ($C\alpha$, $C\beta$, C , $H\alpha$, H , N), each class of 3J scalar coupling constant (ie. $^3J_{HNH\alpha}$ and $^3J_{CC}$) and PREs measured from each nitroxide spin-label as separate experimental *data types*. The experimental datasets used for calculating structural ensembles of A β 40, drkN SH3, ACTR, PaaA2 and α -synuclein

ensembles contain 7, 10, 11, 8, and 14 experimental data types with 237, 458, 608, 541, and 1293 total experimental restraints, respectively.

Calculating experimental observables

All 30 μ s MD simulations analyzed here contained 29976 frames, with a spacing of ~ 1.0 ns per frame.²⁰ Backbone scalar coupling constants were calculated using previously determined Karplus equations.^{57,58} RDCs were calculated with PALES⁵⁹ using a local alignment window of 15 residues.²⁴ Backbone chemical shifts were calculated using SPARTA+¹⁹ with MDTraj.⁶⁰ The SAXS profiles of each frame were calculated using Pepsi-SAXS⁶¹ following a previously described protocol⁶² setting adjustable solvent parameters to $\delta\rho = 3.34 \text{ e/nm}^3$ and $r_0 = 1.68 \text{ \AA}$. The SAXS profiles were scaled by dividing the intensities by the intensities in the first scattering angle in order to normalize the datasets and set the initial intensity to 1. PREs were calculated using distances between $C\alpha$ atoms as described previously⁶³

α -helical order parameter $S\alpha$

$S\alpha$ describes the similarity of each seven-residue fragment of a protein to an ideal helix.³⁰

$S\alpha$ is calculated as follows:

$$S_{\alpha} = \sum_i^N \frac{1 - \left(\frac{\text{RMSD}_{\alpha_i}}{r_0}\right)^8}{1 - \left(\frac{\text{RMSD}_{\alpha_i}}{r_0}\right)^{12}} \quad (12)$$

where RMSD_{α_i} is the root mean square deviation (RMSD) between a seven-residue fragment of a protein (spanning from residue i to residue $i + 6$) and an ideal seven residue helix. Eq. 12 functions as switching function which outputs values ranging from 0 (indicating not helical) to 1 (indicating perfectly helical) for each seven-residue segment. The threshold for this function is adjusted by the parameter r_0 , which we set to 0.80 $\text{\AA}$. Setting the parameter $r_0 = 0.80 \text{ \AA}$ means that a seven-residue segment with $\text{RMSD}_{\alpha_i} > 2.5 \text{ \AA}$ effectively contributes a value 0 to the $S\alpha$ sum, while a seven-residue segment with $\text{RMSD}_{\alpha_i} < 0.5 \text{ \AA}$ contributes a value of 1.0 to the $S\alpha$ sum. The value of $S\alpha$ for a protein conformation can therefore

be interpreted as a proxy for the number of seven-residue fragments that resemble an ideal helix.

Comparison of low-dimensional projections obtained from ELViM and PCA

We utilize the silhouette score⁴ (Eq. 13 and Eq. 14) to compare 2D ELViM and PCA projections in terms of their ability to distinguish between conformational ensembles generated from the 3 force fields investigated in this study: a99SB-*disp*, C22* and C36m. The silhouette score quantifies the consistency of clustering method by measuring how well each cluster is distinguished from it’s nearest neighbors in a lower dimensional projection. For each protein, we perform ELViM and PCA on all three simulation datasets in aggregate such that they are projected onto the same latent space. We consider each "cluster" to consist of all conformations belonging to a given "parent" unbiased MD simulation used for reweighting, and compute the silhouette score accordingly. Consequently, in the following, 'cluster' and 'parent MD simulation ensemble' are synonymous. Distances are measured using the Euclidean distance between projected points. To compute the silhouette score, let $a(i)$ be the average intra-cluster distance for point i in simulation ensemble a , and $b(i)$ be the average inter-cluster distance for point i and all of the points in nearest neighbor simulation ensemble, b . The silhouette score for point i is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (13)$$

The overall silhouette score is the average over all points:

$$\text{silhouette score} = \frac{1}{n} \sum_{i=1}^n s(i) \quad (14)$$

We observe that on average, we obtain substantially large silhouette scores from ELViM projections than PCA Projections (SI Table 1).

Supplementary References

- (1) Teixeira, J. M.; Liu, Z. H.; Namini, A.; Li, J.; Vernon, R. M.; Krzeminski, M.; Shamandy, A. A.; Zhang, O.; Haghighatlari, M.; Yu, L.; others IDPConformerGenerator: a flexible software suite for sampling the conformational space of disordered protein states. *The Journal of Physical Chemistry A* **2022**, *126*, 5985–6003.
- (2) Sterckx, Y. G.; Volkov, A. N.; Vranken, W. F.; Kragelj, J.; Jensen, M. R.; Buts, L.; Garcia-Pino, A.; Jové, T.; Van Melderren, L.; Blackledge, M.; others Small-angle X-ray scattering-and nuclear magnetic resonance-derived conformational ensemble of the highly flexible antitoxin PaaA2. *Structure* **2014**, *22*, 854–865.
- (3) Schwalbe, M.; Ozenne, V.; Bibow, S.; Jaremko, M.; Jaremko, L.; Gajda, M.; Jensen, M. R.; Biernat, J.; Becker, S.; Mandelkow, E.; others Predictive atomic resolution descriptions of intrinsically disordered hTau40 and α -synuclein in solution from NMR and small angle scattering. *Structure* **2014**, *22*, 238–249.
- (4) Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **1987**, *20*, 53–65.
- (5) Cesari, A.; Gil-Ley, A.; Bussi, G. Combining simulations and solution experiments as a paradigm for RNA force field refinement. *Journal of chemical theory and computation* **2016**, *12*, 6192–6200.
- (6) Cesari, A.; Reißer, S.; Bussi, G. Using the maximum entropy principle to combine simulations and solution experiments. *Computation* **2018**, *6*, 15.
- (7) Rangan, R.; Bonomi, M.; Heller, G. T.; Cesari, A.; Bussi, G.; Vendruscolo, M. Determination of structural ensembles of proteins: restraining vs reweighting. *Journal of chemical theory and computation* **2018**, *14*, 6632–6641.

- (8) Pitera, J. W.; Chodera, J. D. On the use of experimental observations to bias simulated ensembles. *Journal of chemical theory and computation* **2012**, *8*, 3445–3451.
- (9) Shannon, C. E. A mathematical theory of communication. *The Bell system technical journal* **1948**, *27*, 379–423.
- (10) Hummer, G.; Köfinger, J. Bayesian ensemble refinement by replica simulations and reweighting. *The Journal of chemical physics* **2015**, *143*.
- (11) Bottaro, S.; Bengtsen, T.; Lindorff-Larsen, K. Integrating molecular simulation and experimental data: a Bayesian/maximum entropy reweighting approach. *Structural bioinformatics: methods and protocols* **2020**, 219–240.
- (12) Köfinger, J.; Stelzl, L. S.; Reuter, K.; Allande, C.; Reichel, K.; Hummer, G. Efficient ensemble refinement by reweighting. *Journal of chemical theory and computation* **2019**, *15*, 3390–3401.
- (13) Köfinger, J.; Hummer, G. Encoding prior knowledge in ensemble refinement. *The Journal of Chemical Physics* **2024**, *160*.
- (14) Bottaro, S.; Bussi, G.; Kennedy, S. D.; Turner, D. H.; Lindorff-Larsen, K. Conformational ensembles of RNA oligonucleotides from integrating NMR and molecular simulations. *Science advances* **2018**, *4*, eaar8521.
- (15) Kummerer, F.; Orioli, S.; Harding-Larsen, D.; Hoffmann, F.; Gavrilov, Y.; Teilum, K.; Lindorff-Larsen, K. Fitting side-chain NMR relaxation data using molecular simulations. *Journal of Chemical Theory and Computation* **2021**, *17*, 5262–5275.
- (16) Flyvbjerg, H.; Petersen, H. G. Error estimates on averages of correlated data. *The Journal of Chemical Physics* **1989**, *91*, 461–466.
- (17) Kish, L. *Survey Sampling*; John Wiley & Sons, Inc.: New York, London, 1965; IX + 643 S., 31 Abb., 56 Tab., Preis 83 s.

- (18) Crehuet, R.; Buigues, P. J.; Salvatella, X.; Lindorff-Larsen, K. Bayesian-maximum-entropy reweighting of IDP ensembles based on NMR chemical shifts. *Entropy* **2019**, *21*, 898.
- (19) Shen, Y.; Bax, A. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *Journal of biomolecular NMR* **2010**, *48*, 13–22.
- (20) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proceedings of the National Academy of Sciences* **2018**, *115*, E4758–E4766.
- (21) Piana, S.; Robustelli, P.; Tan, D.; Chen, S.; Shaw, D. E. Development of a force field for the simulation of single-chain proteins and protein–protein complexes. *Journal of chemical theory and computation* **2020**, *16*, 2494–2507.
- (22) Zhang, O.; Forman-Kay, J. D. Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry* **1995**, *34*, 6784–6794.
- (23) Iesmantavicius, V.; Jensen, M. R.; Ozenne, V.; Blackledge, M.; Poulsen, F. M.; Kjaergaard, M. Modulation of the Intrinsic helix propensity of an intrinsically disordered protein reveals long-range helix–helix interactions. *Journal of the American Chemical Society* **2013**, *135*, 10155–10163.
- (24) Marsh, J. A.; Baker, J. M.; Tollinger, M.; Forman-Kay, J. D. Calculation of residual dipolar couplings from disordered state ensembles using local alignment. *Journal of the American Chemical Society* **2008**, *130*, 7804–7805.
- (25) Ozenne, V.; Schneider, R.; Yao, M.; Huang, J.-r.; Salmon, L.; Zweckstetter, M.; Jensen, M. R.; Blackledge, M. Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *Journal of the American Chemical Society* **2012**, *134*, 15138–15148.

- (26) Wood, G. P.; Rothlisberger, U. Secondary structure assignment of amyloid- β peptide using chemical shifts. *Journal of chemical theory and computation* **2011**, *7*, 1552–1563.
- (27) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: significantly improved protein chemical shift prediction. *Journal of biomolecular NMR* **2011**, *50*, 43–57.
- (28) Lincoff, J.; Haghighatlari, M.; Krzeminski, M.; Teixeira, J. M.; Gomes, G.-N. W.; Gradinaru, C. C.; Forman-Kay, J. D.; Head-Gordon, T. Extended experimental inferential structure determination method in determining the structural ensembles of disordered protein states. *Communications chemistry* **2020**, *3*, 74.
- (29) Bernadó, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *Journal of the American Chemical Society* **2007**, *129*, 5656–5664.
- (30) Pietrucci, F.; Laio, A. A collective variable for the efficient exploration of protein beta-sheet structures: application to SH3 and GB1. *Journal of Chemical Theory and Computation* **2009**, *5*, 2197–2201.
- (31) Ghafouri, H.; Lazar, T.; Del Conte, A.; Tenorio Ku, L. G.; Tompa, P.; Tosatto, S. C.; Monzon, A. M. PED in 2024: improving the community deposition of structural ensembles for intrinsically disordered proteins. *Nucleic acids research* **2024**, *52*, D536–D544.
- (32) Best, R. B.; Zheng, W.; Mittal, J. Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *Journal of chemical theory and computation* **2014**, *10*, 5113–5124.
- (33) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell Jr, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **2017**, *14*, 71–73.

- (34) Tian, C.; Kasavajhala, K.; Belfon, K. A.; Raguetto, L.; Huang, H.; Migués, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; others ff19SB: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *Journal of chemical theory and computation* **2019**, *16*, 528–552.
- (35) Appadurai, R.; Nagesh, J.; Srivastava, A. High resolution ensemble description of metamorphic and intrinsically disordered proteins using an efficient hybrid parallel tempering scheme. *Nature communications* **2021**, *12*, 958.
- (36) Zhang, Y.; Liu, X.; Chen, J. Re-Balancing Replica Exchange with Solute Tempering for Sampling Dynamic Protein Conformations. *Journal of chemical theory and computation* **2023**, *19*, 1602–1614.
- (37) Bussi, G.; Laio, A. Using metadynamics to explore complex free-energy landscapes. *Nature Reviews Physics* **2020**, *2*, 200–212.
- (38) Gong, X.; Zhang, Y.; Chen, J. Advanced sampling methods for multiscale simulation of disordered proteins and dynamic interactions. *Biomolecules* **2021**, *11*, 1416.
- (39) Robustelli, P.; Ibanez-de Opakua, A.; Campbell-Bezatz, C.; Giordanetto, F.; Becker, S.; Zweckstetter, M.; Pan, A. C.; Shaw, D. E. Molecular basis of small-molecule binding to α -synuclein. *Journal of the American Chemical Society* **2022**, *144*, 2501–2510.
- (40) Zhu, J.; Salvatella, X.; Robustelli, P. Small molecules targeting the disordered trans-activation domain of the androgen receptor induce the formation of collapsed helical states. *Nature Communications* **2022**, *13*, 6390.
- (41) Basu, S.; Martínez-Cristóbal, P.; Frigolé-Vivas, M.; Pesarrodoná, M.; Lewis, M.; Szulc, E.; Bañuelos, C. A.; Sánchez-Zarzalejo, C.; Bielskutė, S.; Zhu, J.; others Rational optimization of a transcription factor activation domain inhibitor. *Nature Structural & Molecular Biology* **2023**, *30*, 1958–1969.

- (42) Appadurai, R.; Koneru, J. K.; Bonomi, M.; Robustelli, P.; Srivastava, A. Clustering heterogeneous conformational ensembles of intrinsically disordered proteins with t-distributed stochastic neighbor embedding. *Journal of chemical theory and computation* **2023**, *19*, 4711–4727.
- (43) Heller, G. T.; Aprile, F. A.; Michaels, T. C.; Limbocker, R.; Perni, M.; Ruggeri, F. S.; Mannini, B.; Löhr, T.; Bonomi, M.; Camilloni, C.; others Small-molecule sequestration of amyloid- β as a drug discovery strategy for Alzheimer’s disease. *Science advances* **2020**, *6*, eabb5924.
- (44) Löhr, T.; Kohlhoff, K.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. A small molecule stabilizes the disordered native state of the Alzheimer’s A β Peptide. *ACS Chemical Neuroscience* **2022**, *13*, 1738.
- (45) Bonomi, M.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. Principles of protein structural ensemble determination. *Current opinion in structural biology* **2017**, *42*, 106–116.
- (46) Marsh, J. A.; Forman-Kay, J. D. Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *Journal of molecular biology* **2009**, *391*, 359–374.
- (47) Rao, J. N.; Kim, Y. E.; Park, L. S.; Ulmer, T. S. Effect of pseudorepeat rearrangement on α -synuclein misfolding, vesicle binding, and micelle binding. *Journal of molecular biology* **2009**, *390*, 516–529.
- (48) Hou, L.; Shao, H.; Zhang, Y.; Li, H.; Menon, N. K.; Neuhaus, E. B.; Brewer, J. M.; Byeon, I.-J. L.; Ray, D. G.; Vitek, M. P.; others Solution NMR studies of the A β (1- 40) and A β (1- 42) peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation. *Journal of the American Chemical Society* **2004**, *126*, 1992–2005.
- (49) Bertocini, C. W.; Jung, Y.-S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. Release of long-range tertiary interactions potentiates

- aggregation of natively unstructured α -synuclein. *Proceedings of the National Academy of Sciences* **2005**, *102*, 1430–1435.
- (50) Yan, Y.; McCallum, S. A.; Wang, C. M35 oxidation induces A β 40-like structural and dynamical changes in A β 42. *Journal of the American Chemical Society* **2008**, *130*, 5394–5395.
- (51) Roche, J.; Shen, Y.; Lee, J. H.; Ying, J.; Bax, A. Monomeric A β 1–40 and A β 1–42 peptides in solution adopt very similar Ramachandran map distributions that closely resemble random coil. *Biochemistry* **2016**, *55*, 762–775.
- (52) Lee, J. H.; Li, F.; Grishaev, A.; Bax, A. Quantitative residue-specific protein backbone torsion angle dynamics from concerted measurement of 3 J couplings. *Journal of the American Chemical Society* **2015**, *137*, 1432–1435.
- (53) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. Mapping long-range interactions in α -synuclein using spin-label NMR and ensemble molecular dynamics simulations. *Journal of the American Chemical Society* **2005**, *127*, 476–477.
- (54) Sung, Y.-h.; Eliezer, D. Residual structure, backbone dynamics, and interactions within the synuclein family. *Journal of molecular biology* **2007**, *372*, 689–707.
- (55) Choy, W.-Y.; Mulder, F. A.; Crowhurst, K. A.; Muhandiram, D.; Millett, I. S.; Doniach, S.; Forman-Kay, J. D.; Kay, L. E. Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *Journal of molecular biology* **2002**, *316*, 101–112.
- (56) Kjaergaard, M.; Nørholm, A.-B.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B. Temperature-dependent structural changes in intrinsically disordered proteins: Formation of α -helices or loss of polyproline II? *Protein Science* **2010**, *19*, 1555–1564.

- (57) Li, F.; Lee, J. H.; Grishaev, A.; Ying, J.; Bax, A. High Accuracy of Karplus Equations for Relating Three-Bond J Couplings to Protein Backbone Torsion Angles. *ChemPhysChem* **2015**, *16*, 572–578.
- (58) Vögeli, B.; Ying, J.; Grishaev, A.; Bax, A. Limits on variations in protein backbone dynamics from precise measurements of scalar couplings. *Journal of the American Chemical Society* **2007**, *129*, 9377–9385.
- (59) Zweckstetter, M.; Bax, A. Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *Journal of the American Chemical Society* **2000**, *122*, 3791–3792.
- (60) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophysical journal* **2015**, *109*, 1528–1532.
- (61) Grudinin, S.; Garkavenko, M.; Kazennov, A. Pepsi-SAXS: an adaptive method for rapid and accurate computation of small-angle X-ray scattering profiles. *Acta Crystallographica Section D: Structural Biology* **2017**, *73*, 449–464.
- (62) Pesce, F.; Lindorff-Larsen, K. Refining conformational ensembles of flexible proteins against small-angle x-ray scattering data. *Biophysical journal* **2021**, *120*, 5124–5135.
- (63) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *The journal of physical chemistry B* **2015**, *119*, 5113–5123.